AFFDL-TR-79-3134

AD A080563

PROCEEDINGS OF THE FIFTEENTH ANNUAL CONFERENCE ON MANUAL CONTROL

WRIGHT STATE UNIVERSITY MARCH 20-22, 1979

NOVEMBER 1979

TECHNICAL REPORT AFFDL-TR-79-3134
Interim Report — April 1978 to March 1979

DDC FILE COPY

Approved for public release; distribution unlimited.

AIR FORCE FLIGHT DYNAMICS LABORATORY
AIR FORCE WRIGHT AERONAUTICAL LABORATORIES
AIR FORCE SYSTEMS COMMAND
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

This document contains blank pages that were not filmed.

NOTICE

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely related Government procurement operation, the United States Government thereby incurs no responsibility nor any obligation whatsoever; and the fact that the government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

This report has been reviewed by the Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and is approved for publication.

Frank L George, Editor
Control Dynamics Branch

Vernon O Hoehne
Acting Chief, Control Dynamics Br
Flight Control Division

FOR THE COMMANDER

Morris A Ostgaard
Acting Chief
Flight Control Division

"If your address has changed, if you wish to be removed from our mailing list, or if the addressee is no longer employed by your organization please notify AFFDL/FGC ,W-PAFB, OH 45433 to help us maintain a current mailing list".

Copies of this report should not be returned unless return is required by security considerations, contractual obligations, or notice on a specific document.

AIR FORCE/56780/29 January 1990 — 200

REPRODUCTION QUALITY NOTICE

This document is the best quality available. The copy furnished to DTIC contained pages that may have the following quality problems:

- Pages smaller or larger than normal.
- Pages with background color or light colored printing.
- Pages with small type or poor printing; and or
- Pages with continuous tone material or color photographs.

Due to various output media available these conditions may or may not cause poor legibility in the microfiche or hardcopy output you receive.

☑ If this block is checked, the copy furnished to DTIC contained pages with color printing, that when reproduced in Black and White, may change detail of the original copy.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFFDL-TR-79-3134	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE *(and Subtitle)* Fifteenth Annual Conference on Manual Control		5. TYPE OF REPORT & PERIOD COVERED Interim April 78 - March 79
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Frank L. George, Editor		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Wright State University Dayton, Ohio 45435		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 62201F 99949906
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Flight Dynamics Laboratory (FGC) Wright-Patterson Air Force Base, OH 45433		12. REPORT DATE November 1979
		13. NUMBER OF PAGES 817
14. MONITORING AGENCY NAME & ADDRESS*(if different from Controlling Office)*		15. SECURITY CLASS. *(of this report)* Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT *(of this Report)*
This report has been approved for public release; its distribution is unlimited.

17. DISTRIBUTION STATEMENT *(of the abstract entered in Block 20, if different from Report)*

18. SUPPLEMENTARY NOTES

19. KEY WORDS *(Continue on reverse side if necessary and identify by block number)*

Human Dynamics	Attention Allocation
Human Modeling	Simulation
Man-Machine System	Control Manipulators
Manual Control	Displays
Decision Making	Identification

Control Manual

20. ABSTRACT *(Continue on reverse side if necessary and identify by block number)*
This volume contains the proceedings of the Fifteenth Annual Conference on Manual Control, held at Wright State University, Dayton, Ohio, March 20 - 22, 1979. Fifty-six papers were presented during the conference and all are represented, ~~in this~~ volume, either as complete papers or as abstracts provided by the authors. Topics of the papers include higher order manual tasks involving attention allocation and decision-making, human dynamic modeling and identification, and design methodology for human operator-controlled systems. Also covered were aspects of control manipulators and

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE *(When Data Entered)*

displays, as well as considerations for simulation of human operator-controlled systems.

(8) Interim rept. Apr 78 – Mar 79;

(14) AFFDL-TR-79-3104

15 TH

(6) Proceedings of the

ANNUAL

conference on

MANUAL

control (15th)

Held

MARCH 20 – 22, 1979

(16) 9994

(17) 99

WRIGHT STATE UNIVERSITY, DAYTON, OHIO.

(10) Frank A. Woodyard

(11) Nov 79

(12) 812

SUPPORT PROVIDED BY THE US AIR FORCE FLIGHT DYNAMICS LABORATORY

Accession For
NTIS GRA&I
DDC TAB
Unannounced
Justification
By
Distribution/
Availability Codes
Dist
A

AD A072 070

DC

CONFERENCE CHAIRMAN

Dr William R Wells
Wright State University

AIR FORCE SUPPORT CHAIRMAN

Mr Ronald O Anderson
Air Force Flight Dynamics Laboratory

CONFERENCE ARRANGEMENTS

Mr Frank L George
Air Force Flight Dynamics Laboratory

FOREWORD

This volume, published with the support of the Air Force Flight Dynamics Laboratory, contains the proceedings of the Fifteenth Annual Conference on Manual Control held at Wright State University, Dayton, Ohio from March 20 through 22, 1979. All papers presented at the meeting are represented in this volume. Either complete manuscripts as provided by the authors or abstracts are printed. Both formal and informal papers were presented, as indicated in the Table of Contents.

This was the fifteenth in a series of conferences dating back to December 1964. These earlier meetings and their proceedings are listed below:

First Annual NASA-University Conference on Manual Control, The University of Michigan, December 1964. (Proceedings not printed.)

Second Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, February 28 to March 2, 1966, NASA SP-128.

Third Annual NASA-University Conference on Manual Control, University of Southern California, March 1-3, 1967, NASA SP-144.

Fourth Annual NASA-University Conference on Manual Control, The University of Michigan, March 21-23, 1968, NASA SP-192.

Fifth Annual NASA-University Conference on Manual Control, Massachusetts Institute of Technology, March 27-29, 1969, NASA SP-215.

Sixth Annual Conference on Manual Control, Wright-Patterson AFB, April 7-9, 1970, proceedings published as AFIT/AFFDL Report, no number.

Seventh Annual Conference on Manual Control, University of Southern California, June 2-4, 1971, NASA SP-281.

Eighth Annual Conference on Manual Control, University of Michigan, Ann Arbor, Michigan, May 17-19, 1972, AFFDL-TR-72-92.

Ninth Annual Conference on Manual Control, Massachusetts Institute of Technology, May 23-25, 1973, proceedings published by MIT, no number.

Tenth Annual Conference on Manual Control, Wright-Patterson AFB, April 9-11, 1974, proceedings published as AFIT/AFFDL Report, no number.

Eleventh Annual Conference on Manual Control, NASA-Ames Research Center, May 21-23, 1975, NASA TM X-62,464.

Twelfth Annual Conference on Manual Control, University of Illinois, May 25-27, 1976, NASA TM X-73,170.

Thirteenth Annual Conference on Manual Control, Massachusetts Institute of Technology, June 15-17, 1977, proceedings published by MIT, no number.

Fourteenth Annual Conference on Manual Control, University of Southern California, Los Angeles, CA, April 25-27, 1978, NASA CP-2060.

CONTENTS

		Page
Foreword.....		*iii*
Session 1.	Modeling For Higher Order Tasks Moderator: William H. Levison, Bolt Beranek and Newman	1
1.	*DEMON: A Human Operator Model for Decision Making, Monitoring and Control (Formal)* Ramal Muralidharan and Sheldon Baron	3
2.	*Modeling Human Decision Making in Multitask Situations Involving Both Control and Discrete Tasks (Formal)* T. Govindaraj and William B. Rouse	24
3.	*A Study of Human Attention Allocation in Multiple Process Monitoring Situations (Informal)* Joel S. Greenstein and William B. Rouse	31
4.	*Models for Human Operators Based on Nonuniformly Sampled Systems (Formal)* G.D. Kontopidis, F.H. Glanz, D.E. Limbert	43
5.	*Division of Attention as a Function of the Number of Steps, Visual Shifts, and Memory Load (Formal)* Richard A. Chechile, Keith Butler, William Gutowski, Everett A. Palmer	71
6.	*Attention Allocation in Dynamic Environments (Formal)* Christopher D. Wickens and Pamela Tsang	82
7.	*Can You Control Your Control: On the Efficiency of Attention Allocation Between Tracking Dimensions (Formal)* Daniel Gopher, David Navon	93
8.	*A Discrete Control Analysis of Coordination Activities in a Simulated AAA System (Formal)* R.A. Miller	114
9.	*Progress Report on the Teleconference Cum Electronic Journal on Mental Workload (Informal)* Thomas B. Sheridan	144
Session 2.	Considerations For Man-In-The-Loop Simulations Moderator: Dr. Sheldon Baron, Bolt Beranek and Newman	145
1.	*The Effects of Simulation Fidelity on Air-to-Air Tracking (Formal)* B.R. Ashworth, B.T. McKissick, and R.V. Parrish	147

Page

2. *Effects of Simulator Delays on Performance and Learning in a Roll-Axis Tracking Task (Formal)* 168
William H. Levison, Roy E. Lancraft, Andrew M. Junker

3. *A Performance Analyzer for Identifying Changes in Human Operator Tracking Strategies (Formal)* 187
Greg L. Zacharias and William H. Levison

4. *Optimalisation and Evaluation of Linear Motion Filters (Formal)* 213
R.J.A.W. Hosman, G.A.J. van de Moesdijk, H.C. van der Vaart

5. *Human Pilot Perception Experiments (Informal)* 243
Dr. David L. Quam

6. *Effects of Various Lateral-Beam-Motion Washouts on Pilot Tracking and Opinion in the "LAMAR" Simulator (Formal)* 244
Henry R. Jex, Wayne F. Jewell, Raymond E. Magdaleno

7. *Control Loading Methodology Using a Microcomputer (Formal)* 267
Dr. Gerry Albers

8. *An Approach to Design of Simulation Systems for Training and Selection of Pilots in an Environment of Competition (Formal)* 268
Dr. Leonid Lipchin

Session 3. Human Dynamics Modeling 275
Moderator: Dr. Carroll Day, 6570 Aerospace Medical Research Lab

1. *Effects of Muscle Vibration and Joint Oscillation on Human Motor Mechanisms (Formal)* 277
Gyan C. Agarwal, Gerald L. Gottlieb

2. *Input Adaptive Behavior of the Human Eye Tracking System (Formal)* 300
Loek Kretzschmar

3. *Electrodermal Lability and Capacity for Dual-Task Performance (Informal)* 301
Russell A. Benel, Michael G. H. Coles, Denise C. R. Benel

4. *Predictor Operator in Pursuit and Compensatory Tracking (Informal)* 311
Robert J. Jaeger, Gyan C. Agarwal, Gerald L. Gottlieb

5. *Advances in Modeling Pilot Tracking Performance in the Presence of Sustained Linear Accelerations (Formal)* 331
Jonathan Korn and David L. Kleinman

6. *Investigation of Alternate Human Operator Optimal Control Model Structures (Formal)* 341
Anil V. Phatak

CONTENTS (Continued)

		Page
7.	*A Structural Model of the Adaptive Human Pilot (Formal)* Ronald A. Hess	369
8.	*Fitts' Law and Target Acquisition (Formal)* Richard J. Jagacinski, Daniel W. Repperger, Sharon L. Ward, Betty Glass, Martin S. Moran	380
9.	*Multi-Level and Multi-Task Supervisory Control (Formal)* Dana R. Yoerger, Bahman Daryarian and Thomas B. Sheridan	392
Session 4.	Control And Manipulators Moderator: Dr. Ronald A. Hess, NASA Ames Research Center	393
1.	*How to Talk to a Robot Manipulator (Formal)* Thurston, L. Brooks and Thomas B. Sheridan	395
2.	*A Tactual Aid for the Visual Approach (Formal)* Sheryl L. Chappell, Richard D. Gilson	396
3.	*Evaluation of Proximity Sensor Aided Grasp Control for Shuttle RMS (Formal)* A.K. Bejczy, J.W. Brown, J.L. Lewis	404
4.	*The Transfer of Control and Guidance Information to the Pilot Through the Manipulator Forces (Formal)* K.H. Doetsch and W. Röger	430
5.	*Control/Display System Synthesis Using Closed Loop Performance Criteria and Pilot Modeling (Informal)* John E. Hart	445
6.	*Investigation of a Control Device with Force Input and Displacement Feedback in Steering a Low-Frequency System (Formal)* Hans E. Boller, Walther Krüger	457
7.	*Optimizing Spring Force, Viscous Damping, and Inertia of a Joystick (Formal)* Walther Krüger	471
8.	*Development Testing of a Moveable Sidestick Controller for the Production F-16 (Informal)* Capt Terry L. Lutz	483
Session 5A.	Pilot Model Identification Moderator: Dr. William R. Wells, Wright State University	485
1.	*On the Identification of Parameters in the Optimal Control Model (Formal)* Roy E. Lancraft, David L. Kleinman	487
2.	*Identification of Multivariable Dynamic Behavior of Man by a Correlation Technique (Informal)* B. Tilemann	502

CONTENTS (Continued)

		Page
3.	*Transfer Function Properties of the Optimal Control Model in Target Tracking Tasks (Informal)* A.V. Phatak and K.M. Kessler	515
4.	*Estimation of Parameters Characterizing Human Autonomic Responses to Stress (Informal)* Paul Milgram, John Senders	524
5.	*A Model for Manual Decelerating Approaches to Hover (Informal)* Robert K. Heffley	545
6.	*Predicting Field of View Requirements for VSTOL Aircraft Approach and Landing (Informal)* Warren F. Clement (delivered by R.K. Heffley)	555
Session 5B.	Display Considerations Moderator: Dr. Malcolm L. Ritchie, Wright State University	573
1.	*Extended Event-Driven Displays for Manipulator Control (Formal)* A.K. Bejczy, G. Paine	574
2.	*Further Consideration on Extended Signal Quickening for Manual Control (Informal)* Shigeru Kurosu and Masayoshi Tomizuka	588
3.	*Visual Scene Perception - Frequency Domain Data and Model Parameter Estimation Procedure (Informal)* P.H. Wewerinke	601
4.	*A Study of the Use of Color Graphics for Representing Schematic Information (Informal)* Heidi L. Neubauer and William B. Rouse	624
5.	*Content, Symbology, and Format of Cockpit Display of Traffic Information: General Aviation Pilot Opinion (Formal)* Sandra G. Hart	632
6.	*Perception of Aircraft Separation with Various Symbols on a Cockpit Display of Traffic Information (Informal)* Everett Palmer, Daniel Baty and Sharon O'Connor	650
Session 6.	Design Methodology And Applications Moderator: Mr. Ronald O. Anderson, Air Force Flight Dynamics Lab	667
1.	*An Industrial Utilities Model for Research in Man-Process Control (Formal)* B. Dell Campbell and Richard S. Shirley	669
2.	*The Design and Evaluation of Complex Systems (Formal)* Dennis B. Beringer	690

CONTENTS (Continued)

		Page
3.	*Approaches for Supervisory Management of Information in Automated Airborne Systems (Informal)* **R. Steeb, Y.Y. Chu**	691
4.	*Experiments on Interrupted Monitoring of a Second Order Process (Formal)* **Everett Palmer**	703
5.	*Using Modeling to Determine Measures of Non-Productive Control Activity (Formal)* **D.W. Repperger and A.M. Junker**	725
6.	*Optimal System Augmentation for a Tracking Task with High-Order Dynamics (Formal)* **David K. Schmidt**	749
7.	*An Eclectic Reformulation of Flying Qualities (Formal)* **E.D. Onstott and W.H. Faulkner**	750
8.	*AAA Tracking with Delayed, Discrete-Time Feedback (Informal)* **Arye R. Ephrath**	774
9.	*Analysis of a Tracer Fire System as a Man-Machine Problem (Formal)* **D.W. Repperger and M.M. Vikmanis**	787
10.	*Manual Control Aspects of Tactical Antiaircraft Artillery and Surface to Air Missile Engagements (Informal)* **Capt George J. Valentino**	802
Author Index		803

SESSION 1: MODELING FOR HIGHER ORDER TASKS

Moderator: William H. Levison, Bolt Beranek and Newman

DEMON: A HUMAN OPERATOR MODEL FOR DECISION MAKING, MONITORING AND CONTROL

Ramal Muralidharan and Sheldon Baron

Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, MA

ABSTRACT

This paper describes application of a decision-making, monitoring and control model (DEMON) for the human operator to a task involving control of Remotely Piloted Vehicles (RPVs). The DEMON model is an extension of the Optimal Control Model (OCM) of the operator derived by infusing decision theoretic notions into the basic OCM structure. The resulting model is designed to treat situations in which control actions may be infrequent and monitoring and decision-making are the operator's main tasks.

The analysis of this problem illustrates some of the major considerations in applying DEMON to complex, supervisory control problems. It shows that with fairly straightforward assumptions about the operator's task, DEMON will give reasonable predictions of performance. However, the model results are not compared with actual data so DEMON is presently unvalidated.

1. INTRODUCTION

This paper* describes application of a decision making, monitoring, and control model (DEMON) for the human operator to a Remotely Piloted Vehicle (RPV) en route control task. The task modelled is a simplified version of a simulated RPV mission [1,2]. It retains many of the cognitive aspects of the full simulation but differs in several details, particularly with respect to the operator/system interface. Thus, it allows us to develop an understanding of the modelling approach without undue complication.

The DEMON model is an example of the so-called top-down or analytic approach to human performance modelling. Such an approach begins with a mathematical characterization of the task including the overall goals and the criteria for good performance. Then, one attempts to develop the assumptions about the human operator and the system that are necessary and sufficient to

* The research reported here was supported by AFOSR-AMRL under contract F44620-76-C-0029.

characterize performance in relation to the parameters of interest to system designers.

DEMON does, in fact, have its foundation in control theory and in statistical estimation and decision theory. In particular, it draws heavily on the information processing model implicit in the optimal control (OCM) model of the human operator (see, e.g., [3] for a recent review of the OCM). To this information processing structure is added a decision-making structure for modelling discrete monitoring and control decisions and a structure for computing continuous control actions.

The decision making structure in DEMON embodies the concept of expected net gain (ENG), which is used as a criterion for making a rational choice among alternatives. The expected net gain ENG from a particular action is obtained by subtracting the cost of that action from its expected gain. The expected gain itself is the difference between the expected cost of events when no action is taken and the expected cost of events that may arise after this action. The rational choice is to select that action which has the greatest ENG.

The DEMON modelling approach views the human (en route) operator as an element in a closed-loop control system, as shown in the block diagram of Figure 1. The elements in this diagram are described in [4].

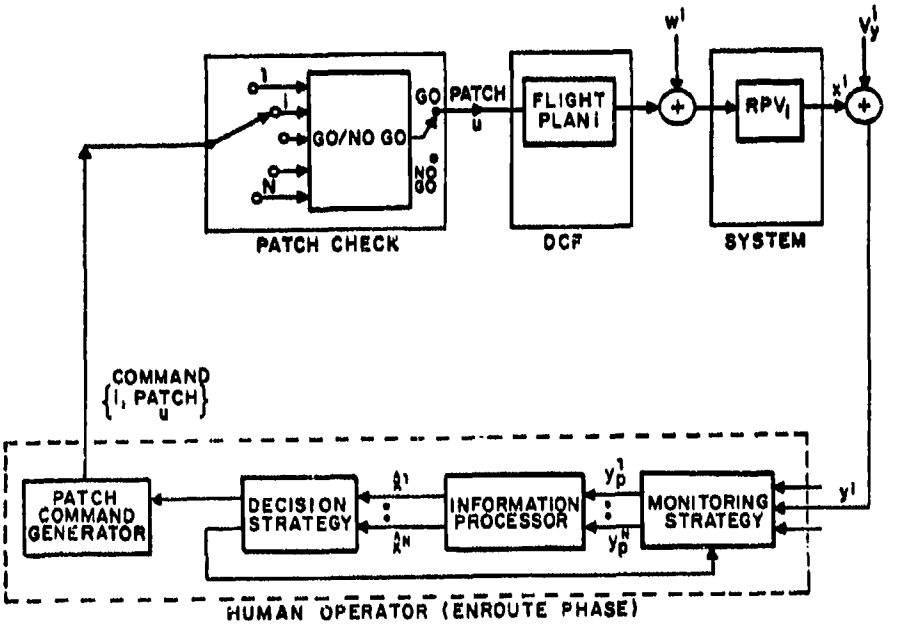


Fig. 1 DEMON View of RPV Control Problem

2. APPLICATION OF 'DEMON' TO A SIMPLIFIED RPV CONTROL PROBLEM

2.1 System Description

In developing and demonstrating the DEMON model in the RPV context we decided that it would be most efficient to consider a simplified version of the problem. The version used in this study is discussed below. A more complete and general system model for DEMON application is given in [2,4].

2.1.1 Flight Plan

RPV's are launched as triads consisting of strike (S), electronic countermeasure (E), and reconnaissance (L) vehicles. Nominal flight plans for each triad of RPVs are shown in Figure 2. The S-vehicle is launched first followed by an E-vehicle and an L-vehicle forty-five seconds apart. The flight plans are such that the E-vehicle arrives in the terminal area first, fifteen seconds ahead of the S-vehicle and 90 seconds ahead of the L-vehicle. All of the flight plans contain preprogrammed turns as shown. The duration of the en route phase of the mission is about 1200 seconds.

In most of our simulations there will be six RPVs, two triads, to control during the course of the mission. The second triad will be launched sequentially as above, 135 seconds after launch of the first triad. Nominal pop-up and hand-off times for the six RPVs are selected to achieve the mission objectives with the given flight plan. The nominal flight plan was selected so that all RPVs exceed the tolerable LATDEV errors of 1500 ft. Moreover, if uncorrected, ETA errors for RPVs 1,2,4, and 6 would also exceed the 15 sec (1.0125×10^4 ft.) tolerance required for proper sequencing.

2.1.2 State Equations

We wish to describe the manner in which RPVs deviate from the nominal flight plans and how they are controlled so as to keep these deviations small. A description that is completely faithful to the full simulation requires four state variables per RPV as shown in Figure 3 and, given present program constraints, seriously limits the number of RPVs that can be considered in a single DEMON simulation.* We, therefore, chose a simpler model for describing deviations from nominal. In Figure 3 let B denote the desired RPV position and P the actual position at time t so that

x_1 = ground speed (ETA) error in feet

x_2 = LATDEV error in feet

---

* One can, after some manipulation reduce the number of states to three by considering equations for speed, cross-track and heading errors.

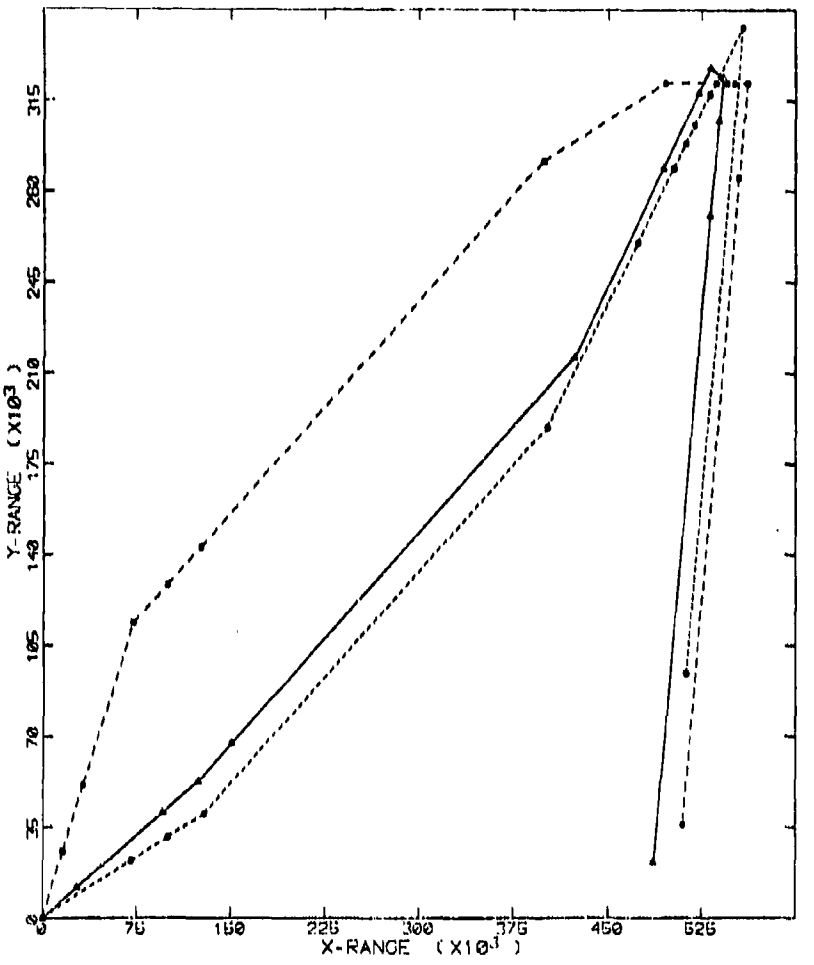


Fig. 2 Flight Plan for a Triad of RPVs

Then, we set

$$\dot{x}_1(t) = u_1(t) + w_1(t) \quad (1)$$
$$\dot{x}_2(t) = u_2(t) + w_2(t) \quad (2)$$

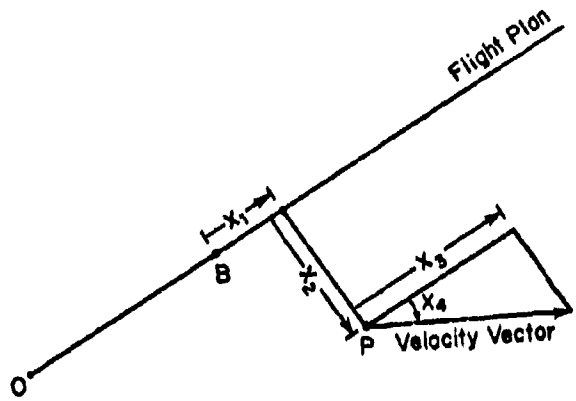


Fig. 3 Choice of Co-ordinates for System Equation

where $u_1(t)$ and $u_2(t)$ are control inputs to be selected by the operator and $w_1(t)$ and $w_2(t)$ are disturbances causing flight plan errors.

The control variables are defined in a manner which retains some of the coupling between ETA errors:

$$\begin{aligned} u_1(t) &= 0 && \text{if no patch is in effect} \\ &= V\cos\phi(t) - \bar{V} && \text{otherwise} \end{aligned} \qquad (3)$$

$$\begin{aligned} u_2(t) &= 0 && \text{if no LATDEV patch is in effect} \\ &= V\sin\phi(t) && \text{otherwise} \end{aligned}$$

where V is the true airspeed, $\bar{V} = 675$ ft/sec is the nominal speed and $\phi(t)$ is a pseudo "heading" given by

$$\phi(t) = \tan^{-1}\left(x_2(t)/x_1(t)\right) \qquad (4)$$

Equations (1)-(3) are shown for a single RPV. Similar equations govern the flight of the other RPVs. The operator's patch commands then correspond to selecting a change in velocity or heading so as to eliminate perceived errors. The possible changes are limited by vehicle constraints on speed, viz.

$$420 \text{ ft/sec} = V_{min} < V < V_{max} = 800 \text{ ft/sec} \qquad (5)$$

and on turn radius, $R_{min} = 5280$ ft.

The disturbances are constant over any leg of the flight plan, but a new constant is chosen after each turn. The values for these constants were chosen randomly to be

w_1 = -6 ft/sec or + .3 ft/sec with equal probability

w_2 = $\pm$ 8 ft/sec with equal probability (6)

in each leg for our basic investigations. Once the pseudo-random sequence of errors was selected it was kept fixed for all remaining parametric investigations.

2.1.3 Displays

We assume a very simple display configuration. A menu display (M) provides ground speed error and a status display (S) provides LATDEV error. On either display the information is available for only one RPV at a time. Thus, in any frame, the operator can only determine either ground speed error or LATDEV error for a single RPV.

We also assume that there are reporting errors associated with the displayed quantities. Thus, the menu and status displays provide, for the selected RPV, noisy information on the states:

MENU: $y_1 = x_1 + v_{y_1}$

STATUS: $y_2 = x_2 + v_{y_2}$ (7)

The reporting errors v_{y_1} and v_{y_2} are assumed to be white noise sequences with an autocovariance that scales with the mean-squared level of the measurement.

2.2 Mathematical Details of the DEMON RPV Operator Model

The essence of the top-down approach is to characterize the mission goals and criteria for good performance in a manner that allows one to predict operator strategies and overall system performance. In the DEMON model, this implies selecting criteria that can be translated into appropriate monitoring and control (patching) strategies. In this section we discuss these criteria and strategies, which are the prime concern of this study, in relation to the problem described above. Other elements of the DEMON operator model shown in Figure 1 are also discussed with respect to the specific application being considered.

2.2.1 Information Processor

The information processing portion of DEMON is drawn directly from the OCM which has been documented extensively [3].

The estimate $\hat{x}$ is the conditional mean of x based on the observations. This quantity and the variance of the estimation error provide a sufficient statistic for specifying the **subjective** probability distribution of x, given the assumptions we have made. It can be shown that patching (u) affects the mean of the subjective distribution whereas monitoring affects both the mean and the uncertainty.

The initial state estimate and uncertainty are parameters of the model that describe the operator's initial knowledge of the state of the system. These parameters become less significant with increasing time. The covariance WOM is also a parameter. It describes the rate at which the operator's uncertainty grows with time in the absence of any additional information. This parameter relates to the operator's expectations concerning the disturbances perturbing the path as determined from instructions or through training.

2.2.2 Monitoring Strategy

Prior to a frame update when new information becomes available, the operator must decide which display to monitor (and perhaps act upon) in the next frame. Because there are two separate displays for ETA and LATDEV for each of the N RPVs and because the operator may choose to do something else,* there are 2N+1 alternatives among which to choose. We assume the operator's choice is a rational one governed by his expectations as to the system behavior, his goals and priorities, and his instructions. For example, performance priorities for the operator to achieve mission objectives in the full RPV study [1] were: i) pop-up (down) in good time sequence, ii) minimize ETA error at hand-off, iii) time phased RPV arrivals, iv) minimize lateral deviation errors, v) minimize command traffic (allow for possible "jamming"), and vi) minimize missed strikes.

The mission factors may be incorporated in an expected net gain criterion of the form

$$ENGM(i) = C_i P_i - CM_i \quad ; \; i=1,2,\ldots,2N$$

Here the constant C_i is the cost associated with the event that the state x_i exceeds some "tolerable" threshold TM_i, while P_i is the **subjective** probability of this event. As stated earlier, the information processor provides the subjective probability distribution for the states x_i of the RPV system under consideration, and, hence, the P_i.

---

* For example, monitor other RPV's not accounted for in the basic state space model in DEMON, such as those on the return list.

The constant CM_i is the action cost of monitoring the display y_i. Assuming that only a single displayed variable is selected for monitoring at a given time, the operator would monitor that y_i which has the greatest (positive) ENGM(i) associated with it. If none of the ENGM(i) is positive, no y_i will be monitored at that time which corresponds to the operator doing "other things".

The parameters in the expression for the expected net gain from monitoring are the threshold, TM_i, associated with the P_i, and the costs C_i and CM_i. These are the quantities that reflect mission objectives, etc.

There is considerable flexibility and some redundancy with respect to choices among these parameters. In particular, there is no requirement that they be constant in time. Thus, we can allow the monitoring thresholds to shrink linearly (or nonlinearly) with time to reflect the fact that a deviation of a given magnitude is much more important near pop-up than it is at launch. Alternatively, the threshold could be constant, reflecting an equal concern for errors throughout the mission. The relative magnitudes of C_i (or CM_i) may be chosen to emphasize the relative importance of ETA over LATDEV or of one RPV (say a strike) over another.

The CM_i can also be used to account for the importance or priority associated with pop-up or hand-off. This is accomplished by letting $CM_i = CM_{io} - CMP_i - CMH_i$. Here, CM_{io} is a constant, reflecting some basic monitoring cost that is fixed during the en route phase. CMP_i is chosen to be zero from launch until τ_p seconds prior to the scheduled pop-up time T_p at which time it takes on a positive value. This value of CMP_i is kept constant until pop-up is completed. Similarly CMH_i has a value of zero until τ_H seconds prior to scheduled hand-off at which time it is given a positive value which is maintained until hand-off is complete. Note that the quantities τ_p and τ_H reflect operator preferences and/or instructions with respect to pop-up and hand-off.

2.2.3 Patching Strategy

By a patching strategy we mean the rational decision whether or not to issue a pop-up, hand-off or a patch (ETA or LATDEV) command. The patching strategy will depend on the monitoring strategy. For example, in the present implementation we assume that monitoring ETA is essential for making a velocity patch and that monitoring LATDEV is essential for making a LATDEV patch. It is assumed that pop-up or hand-off of an RPV may be done while monitoring either of its associated displays.

Just as for monitoring, we assume that the decision to patch is made on the basis of criterion functions that reflect mission goals, etc. Given that the operator is looking at the status display, the choice of a LATDEV patch, pop-up or hand-off command will depend on the expected net gain associated with these commands. We define the expected net gain associated with each of his possible patch actions by the following:

$$ENGL(i) = CL_i(\hat{x}_{2i})^2 - CLK_i$$

$$ENGP(i) = CP_i \exp(t-T_p+\tau_p) \quad [\exp(2t_F)-1]$$

$$ENGH(i) = CH_i \exp(t-T_H+\tau_H) \quad [\exp(2t_F)-1]$$

The rationale for selecting the LATDEV patching criterion ENGL(i) is as follows: $CL_i(x_{2i})^2$ is a measure of the cost of the lateral deviation x_{2i} of the i-th RPV from its flight plan. The purpose of a LATDEV patch is to reduce the lateral deviation x_{2i} to zero. It can be shown that, given the estimate $\hat{x}_{2i}$ and its uncertainty, the operator's expected gain from patching is simply $CL_i(\hat{x}_{2i}^2)$. As in the case for monitoring, this potential gain must be balanced against a patching cost and this is reflected in the CLK_i. A reasonable choice seems to be to have the CLK_i decrease with time. Rewriting $ENGL(i) = CL_i[\hat{x}_{2i})^2 - TL_f^2]$, where $TL_f^2 = CLK_i/CL_i$, we may use the LATDEV objective of the mission to select the value TL_i at hand-off. The value of TL_i at launch is then a parameter to be selected.

Turning now to ENGP(i), the quantities T_p and τ_p were explained earlier while discussing the monitoring strategy. Note that the expression $[\exp(2t_F)-1]$ is included to account for the two-frame ($2t_F$) duration required for a pop-up. $CP_i \exp(t-T_p+\tau_p)$ is the assumed cost of missing a pop-up where the constant CP_i behaves similarly to CMP_i discussed earlier. That is, it has a value zero until τ_p seconds prior to pop-up at which time it takes on a positive constant value until pop-up is completed. Although this positive constant may be used as a parameter of the model, we may also choose CP_i by rationalizing the cost of missing a pop-up against that of losing an RPV because of excessive LATDEV, i.e. by setting the costs equal and solving for CP_i. Of course, the same arguments apply to the ENGH(i).

The identical problem faces the operator when viewing the menu display; only here the ENG from a velocity patch must be compared with ENGP(i) and ENGH(i). The expected net gain for a velocity patch is defined analogously to that for a LATDEV patch, namely by

$$ENGV(i) = CV_i\ (\hat{x}_{2i-1}/V_i)^2 - CVK_i = CV_i[(\hat{x}_{2i-1}/V_i)^2 - TL_f^2$$

with the choice of CV_i and CVK_i governed by similar considerations.

In summary, while monitoring LATDEV, based on his expectations the human operator (in the DEMON implementation) will decide to pop-up, hand-off or do a LATDEV patch depending on which of the expected net gains is most positive. If none of them is positive he will not patch, pop-up, or hand-off. Note that his decision not to patch may depend on his concern for breaking radio silence (as reflected in the shrinking "threshold" TL_i). Similarly, while monitoring ETA he may decide to pop-up, hand-off, do a velocity patch or preserve radio silence.

2.2.4 Patch Command Generator

Once a decision is made to patch a particular RPV, it is necessary to compute and execute the patch control. The purpose of a patch control is to guide the RPV from its current location and heading to intercept and fly along the planned flight path. We assume a simple strategy of minimizing the time to return to the planned flight path assuming that the two control actions on a given RPV are non-interactive. In this case the control actions are trivially computable using the current estimate of the ETA and LATDEV states:

$$u_i = \hat{x}_i/T$$

where T is the duration over which the patch is to take place. To relate the u_i to the velocity and heading we recall from Section 2.1 that

$$\begin{aligned} u_{2i-1} &= V_i \cos \phi_i - \bar{V}_i \\ u_{2i} &= V_i \sin \phi_i \end{aligned}$$

which shows the interaction of the velocity patch V_i and the LATDEV patch ϕ_i. Note that in the two-state formulation of the RPV problem there is no true heading. However, we use the above equations to define "pseudo constant headings" ϕ_i during the LATDEV patch.

A check is made on V_i to see of it is within the allowable limits. If it exceeds the limit then V_i is reset at the limit by adjusting the patch time T and hence the patch control action u_{2i-1}. Then u_{2i} is adjusted to account for its interaction with the velocity patch.

A LATDEV patch will be computed as a constant pseudo heading

$$\phi_i = \text{ARCSIN}\ (u_{2i}/V_i)$$

where

$$u_{2i} = -\hat{x}_{2i}/T$$

A "pseudo radius of turn" $R = 0.5\ V_i T/\phi_i$ is computed .If R violates the specified minimum turn radius RMIN for the RPV then the patch time T is relaxed to satisfy the turn radius constraint. To reflect the operator's experience on making LATDEV patches, a factor called SFACTR is introduced. SFACTR ranges from .2 to 1 and is nominally 1. It is decremented in steps of .2 for each consecutive disallowed LATDEV patch and incremented in steps of .1 for each consecutive successful LATDEV patch. The operator will use SFACTR as a safety factor to avoid tight turns for the LATDEV patch and will use RMIN/SFACTR as a guide to select the radius of turn for the LATDEV patches. Having decided on a patch time T, the control action u_{2i} is recomputed and then u_{2i-1} adjusted to reflect the effect of LATDEV on u_{2i-1}.

3. RESULTS

We now examine the sensitivity of the performance predicted by DEMON to changes in parameters of the system and in those describing operator behavior.

3.1 Basic Parameter Values

The parameters of interest are exhibited in Table 1 along with a set of "basic" values for them. The basic parameter values were chosen as indicated in [5].

Several DEMON runs were made to ascertain the sensitivity to various model parameters. These sensitivity results are described below.

Subsequent results are intended to show overall trends of the effects of parameter changes. These are obtained from a set of summary statistics output by the implemented DEMON model.

3.2 Monitoring Performance

The DEMON model was exercised on several selected values for the parameters characterizing the human operator. Figure 4 shows the effect of varying the monitoring cost CM_i on the monitoring frequency. Two curves are shown. For the curve labelled TM = 1:1, the monitoring thresholds TM_i are constant from launch to hand-off. For the curve labelled TM = 4:1 the monitoring thresholds at launch are four times what they are at hand-off and decrease linearly with mission time. It is seen from Figure 4 that the effect of increasing monitoring cost is to decrease the monitoring frequency (averaged over all the 6 RPVs under consideration). For any given monitoring cost CM, widening the monitoring threshold also decreases the monitoring frequency.

Figure 5 shows the effect on the TM = 1:1 curve in Figure 4 if the operator assumes a different model WOM for the navigation errors. Recall that the ENGM expression (with $C_i=1$) depends on the difference P_i-CM_i. The subjective probabilities P_i depend both on the estimates $\hat{x}_i$ and the uncertainties σ_i associated with these estimates. A lower value of WOM causes lower rate of growth of σ_i and hence it takes longer for the P_i associated with different RPVs to be 'equalized' and start overwhelming the monitoring cost CM_i. Thus for the low WOM case, the effect of CM_i is more pronounced than in the high WOM case. In fact, it can be seen from Figure 5 that for any fixed monitoring cost CM, the monitoring frequency is higher if the operator expects the navigation errors to be greater (as represented by the higher noise covariance WOM). The lower value of WOM also reduces the effect of the pop-up cost, CMP_i. This results in missing scheduled pop-up times as indicated in Table 2 (presented later).

The effect of adding more RPVs on the operator's monitoring behavior is illustrated in Figure 6. In this simple exercise of the DEMON model, the RPVs

Table 1 Important Parameters in the DEMON Model

ID	SYMBOL	DESCRIPTION	VALUES FOR BASIC CASE
	SYSTEM	CHARACTERISTICS	
S1	N	Number of RPV's (RPV "density")	6
S2	R_{MIN}	Minimum allowable turn radius	5280 ft.
S3	t_F	Display update rate	5 sec.
S4	FLTPLN	Flight Plan Parameters such as flight duration and desired time for launch, turns, pop-up, hand-off, turn angles, etc.	Figure 1
S5	w_i	Navigation errors	Eq. (3)
S6	V_{y_i}	Reporting error	-20 dB
HUMAN OPERATOR CHARACTERISTICS			
H1	TM_i	Monitoring threshold	250 ft.
H2	C_i	Cost of exceeding TM_i	1
H3	CM_i	Monitoring Action Cost	0
H4	CL_i	Cost factor for LATDEV	1
H5	TL_i	LATDEV Patch cost	$(250)^2$
H6	CV_i	Cost factor for ETA deviation	109375
H7	TV_i	Velocity patch cost	$(3375)^2$
H8	CP_i	Pop-up cost factor	$2.19E7/Exp(\tau_p)$
H9	τ_i	Operator preferred pop-up interval	5 sec.
H10	CH_i	Hand-off cost factor	CP_i
H11	τ_H	Operator preferred hand-off interval	5 sec.
H12	WØM	Operator's understanding of system navigation errors	101250 10000

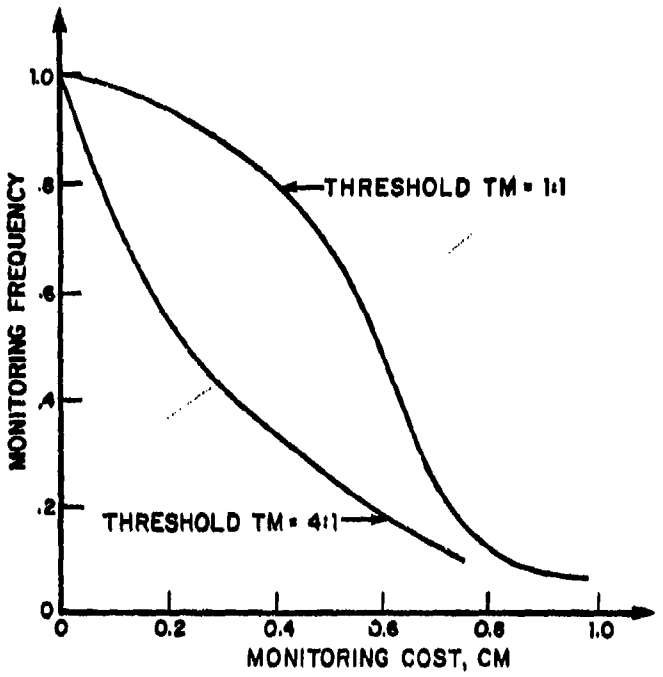


Fig. 4 Effect of Monitoring Cost CM (and Threshold TM) on Monitoring Frequency

in the operator's "en route list" were increased from 1 to 6 while the pseudo random navigation error sequence driving RPV 1 remained unaltered. The monitoring looks on the S and M displays of RPV 1 are plotted against "workload" of the operator as measured by the number of RPVs in his list. Clearly, when the other RPVs compete for attention, the monitoring activity on RPV 1 decreases.

The present implementation of the DEMON model can only accommodate a maximum of 7 RPVs. To investigate the effect of a higher number of RPVs we can introduce the concept of an "equivalent RPV". We saw in Figure 6, that additional RPVs compete for monitoring attention. Likewise in Figures 4 and 5, the effect of increasing the monitoring cost CM is to decrease the monitoring frequency. Recall that CM can be used to account for the necessity of doing other things. High enough values of CM will result in periods in which the six basic RPVs are unmonitored. If we assume that these periods are devoted to other RPVs, and that all RPVs share monitoring attention approximately equally, then an "equivalent" number of RPVs is obtained by dividing the actual number

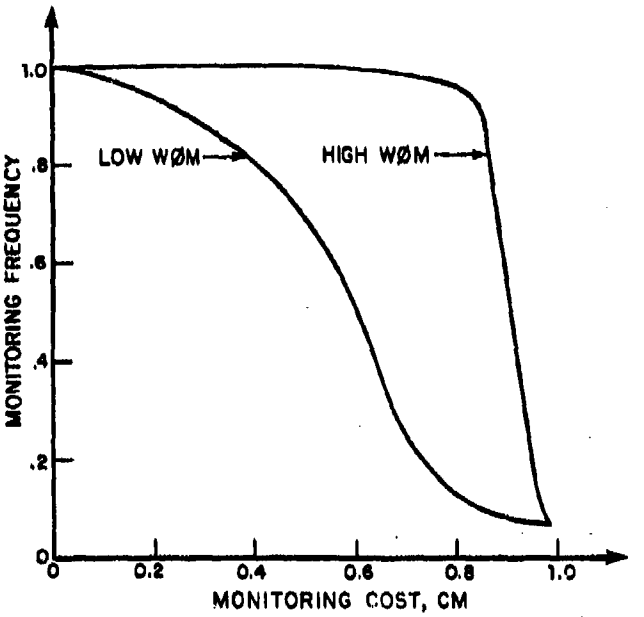


Fig. 5 Effect of Pseudo Process Noise WQM on Monitoring Frequency

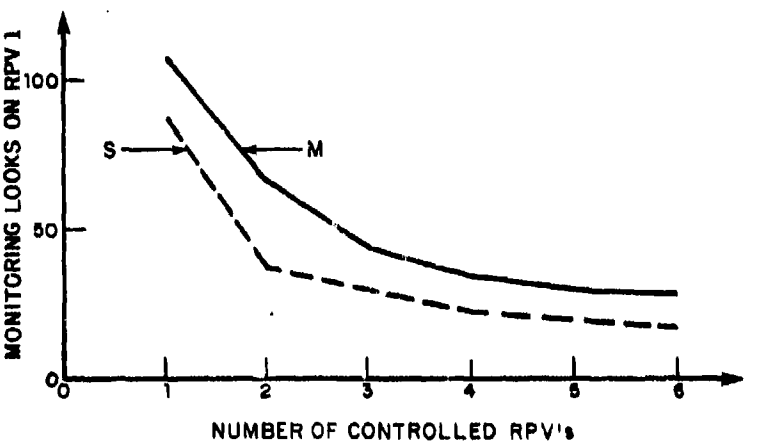


Fig. 6 Effect of Workload (Number of Controlled) RPVs) on Monitoring Looks on RPV 1

under control by the fraction of the total time that they are being observed. For example, if a value of CM is chosen such that the basic six RPVs are monitored only 60 per cent of the time, then we can interpret this as equivalent to ten RPVs under control.

Using the concept of "equivalent RPVs" Figures 4 and 6 are combined to obtain the effect of workload (equivalent over the number of real RPVs in the operator's list). The result plotted in Figure 7 shows that average monitoring frequency is roughly proportional to the reciprocal of the equivalent number of RPVs.

3.3 Patching Performance

In Section 3.2, we considered the sensitivity of monitoring performance to the operator parameters CM, TM, WOM and the system parameter N. We now turn to the study of the operator's effectiveness in controlling the system performance.

The effect of varying patch costs is illustrated in the computer generated graphs of ETA deviation for RPV 1. Figure 8(a) is for the case of flat TL_i, TV_i and Figure 8(b) is for the case of TL_i shrinking to a value at hand off the same as that for the flat TL_i, starting from a value at launch equal to four times that at hand-off. Clearly, higher patch costs during the early part of the mission inhibit patches and result in greater ETA deviations. However, the shrinking TL_i ensures sufficient patching activity to keep ETA near hand-off to be comparable to that in the flat TL_i case.

Figure 9 shows the effect of workload (in terms of equivalent number of RPVs) on the normalized controlled RPV performance as defined below. RPV performance is measured in terms of the RMS values of the deviations (ETADEV and LATDEV) from the flight plan. The RMS values are obtained by "time averaging" and thus include errors all along the flight plan from launch up to hand-off. A controlled RPV is a real RPV on the operator's en route list. The RMS value was averaged over the controlled RPVs and then divided by the indifference threshold setting (chosen to be 250' for LATDEV and 5 sec. for ETADEV in the model run) to obtain the normalized controlled RPV performance. The results for LATDEV and ETADEV are plotted on Figure 10. The operator is able to control up to four RPVs without a degradation in performance. Then, performance steadily deteriorates as workload increases. It is clear that for the chosen parameter settings there is a critical point at N=7. The average RMS LATDEV, integrated from launch to hand-off for N=7, has a value of roughly 1500' with performance deteriorating at a high rate.

It is interesting to note that the RMS errors seem to flatten out for N>15. A further analysis shows that these peak values of RMS errors are the errors inherent in the mission due to navigation errors, that is, the errors that would result if the operator's en route list were "empty" and the RPVs try to negotiate the flight plan without control.

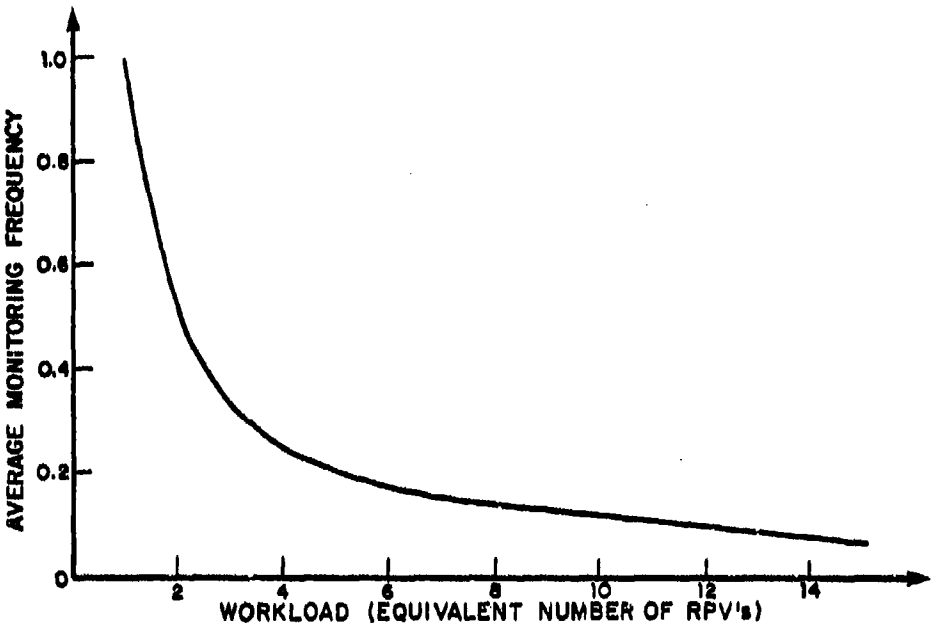


Fig. 7 Effect of Workload (Equivalent Number of RPVs) on Average Monitoring Frequency

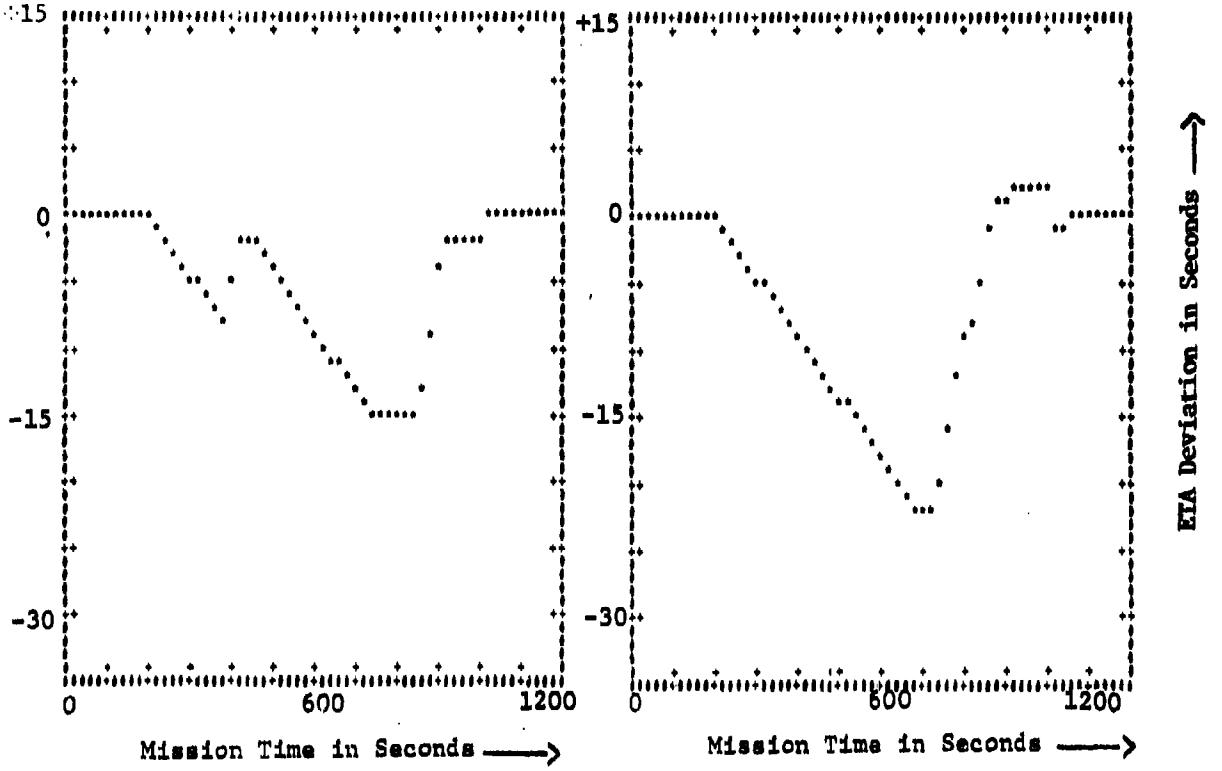


Fig. 8 a) The Effect of Varying Patch Costs on the ETA State Trajectory of RPV 1

Cost at Launch = Cost at Hand-off

Fig. 8 b) The Effect of Varying Patch Costs on the ETA State Trajectory of RPV 1

Cost at Launch = 2 Times Cost at Hand-off

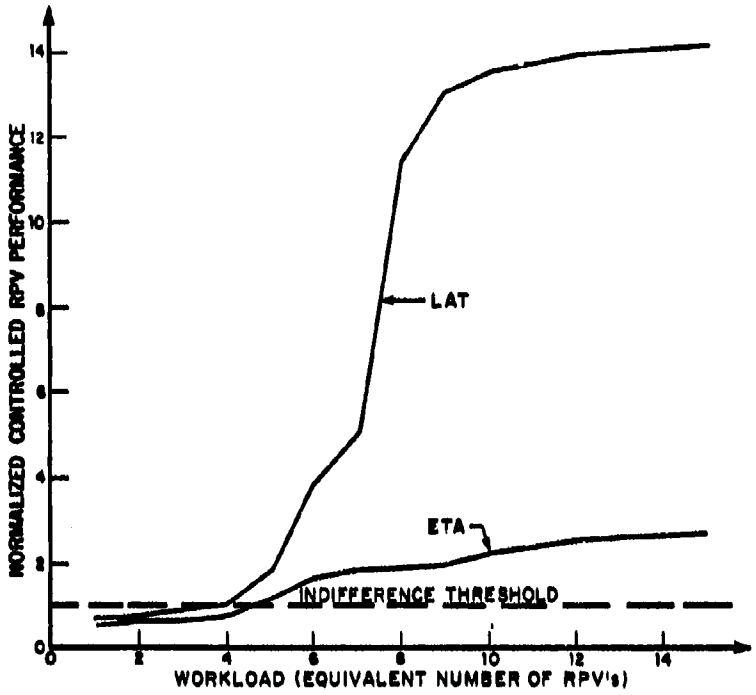


Fig. 9 Effect of Workload (Equivalent Number of Controlled PRVs) on Normalized Controlled Performance for Low WOM.

Table 2 Error is Pop-up Time

RPV #	1	2	3	4	5	6	Average
$\tau_p = 5$	55	10	5	0	5	0	12.5
$\tau_p = 2$	55	10	5	0	5	0	12.5
$\tau_p = 0$	50	10	5	5	40	5	19.1

Now we turn to the effect of τ_p on pop-up. Recall that τ_p seconds prior to the desired pop-up time T_p, the DEMON operator begins to be concerned about pop-up. A few model runs were made by varying τ_p and the results are displayed in Table 2. As it was mentioned earlier the lower WOM causes the DEMON operator to miss pop-ups even with a τ_p of 5 sec. It is seen from Table 2 that with τ_p=0 the average error in pop-up time is even larger. These errors in pop-up time also cause similar errors in hand-off time which increases from an average value of 12.5 for τ_p=5 to 31.7 for τ_p=0. Note that the 5 sec. frame update time makes all intermediate values of τ_p between 0 and 5 behave the same as τ_p = 5 sec.

The effects of various human parameters on monitoring and patching performance were discussed above. The only system parameter that was included in the above discussion was the workload as represented by the number N of RPVs for which the operator has responsibility. We consider now the effect of the reporting error V_y and the navigation errors w_i on the system performance. Table 3 shows the effect of various levels of reporting errors V_y on the RMS deviations (averaged over the six controlled RPVs). It is clear that higher reporting errors lead to degradation in performance. The mean error in pop-up time is also shown. We would expect the errors in pop-up time to degrade steadily with increasing reporting errors if these errors were obtained by averaging over many runs (i.e., by Monte Carlo simulation).

The effect of navigation errors is shown in Table 4 for three cases. Case A is the basic case, Case B has higher ETA navigation errors over Case A and Case C has higher LATDEV navigation errors over Case B. The results again show degradation in RMS deviations as the navigation errors get worse. Table 4 also shows that average number of patches per RPV also increases. The mean error in pop-up time is also displayed in Table 4 and the discussion in the previous paragraph applies here as well.

4. CONCLUSION

In summary, the results obtained with the DEMON model are very interesting and are representative of the type of results that may be obtained with a top-down approach to modelling the RPV control problem. Of course, more sensitivity results would be useful but the results obtained thus far for monitoring performance show that the model does behave reasonably, that the parameters do significantly affect the performance and that the monitoring and patching trends are as expected. The parameters seem to be sufficient to capture the important aspects of variations in monitoring and patching strategies. The results also show how the model may address important considerations from the system designer's point of view such as RPV/Operator ratio, allowable navigation errors, tolerable reporting errors and so forth. However, in the present version of DEMON details of control and display implementations are not so readily addressed, although these may be dealt with by further refinements of the model.

Table 3: Effect of Reporting Errors

Vy	-20	-17	-14
RMS ETADEV (1000 ft.)	5.57	5.94	6.33
RMS ETADEV (1000 ft.)	0.95	1.02	1.15
MEAN ERROR in pop up time (sec.)	12.5	8	20

Table 4: Effect of Navigation Errors

	CASE A	CASE B	CASE C
w_1 ft/sec w_2 ft/sec	0.6 0.8	6.7 0.8	6.7 2.0
RMS ETA DEV (1000 ft.)	5.57	9.13	10.38
RMS LAT DEV (1000 ft.)	0.95	1.75	2.76
MEAN ERROR in pop up time (sec.)	12.5	24.2	22.5
NUMBER OF PATCHES (ETA + LAT DEV)	130	146	152

Thus, the results obtained do appear to provide a 'proof of concept' for the top-down approach employed in DEMON for modelling the RPV en route control problem. However, validation of the model requires comparison of the results from DEMON with data from a 'real' RPV mission, which remains to be done.

REFERENCES

1. Mills, R.,R. Bachert and N. Aume, "Supplementary Report of the RPV System Simulation Study II : Evaluation of RPV Position Report Smoothing and Automatic Heading Correction", AMRL-TR 75-87, September 1975.

2. Miller, D., Feehrer, C., Muralidharan, R., Pew, R. and Baron, S., "Development of Human Performance models for Man-Machine System Simulation", BBN Report No. 3739, October 1978.

3. Baron, S., "A Model for Human Control and Monitoring Based on Modern Control Theory," **Journal of Cybernetics and Information Sciences**, Vol. 4, No. 1, Spring 1976.

4. Muralidharan, R. and S. Baron, "Combined Monitoring Decision and Control Model for the Human Operator in a Command and Control Task," Proceedings of the Fourteenth Annual Conference on Manual Control, NASA Conference Publication, 2060.

5. Muralidharan, R.,S. Baron and C. E. Feehrer , "A Decision, Monitoring, and Control Model of the Human Operator Applied to an RPV Control Problem," BBN Report No. 4075, March 1979.

MODELING HUMAN DECISION MAKING IN MULTITASK SITUATIONS INVOLVING BOTH CONTROL AND DISCRETE TASKS*

T. Govindaraj and William B. Rouse
Department of Mechanical and Industrial Engineering
Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801

ABSTRACT

In many multitask situations, the human operator must allocate his attention between control and discrete tasks so as to optimize some measure of total performance. When preview of the reference trajectory is available, scheduling of control and discrete tasks is possible. Control activity is intermittent to accommodate discrete tasks at various times. A model has been developed for this situation using dynamic programming to solve an optimal control problem. Intermittent control activity is achieved by suitably varying the weights on control along the planning horizon. This has the desired feature of redistributing the control activity to result in optimum performance. An experiment has been designed where subjects control an airplane symbol over a map, shown a fixed distance into the future. Discrete tasks made up of number entry tasks are introduced with different arrival rates. Experiments in progress are described.

INTRODUCTION

For quite some time now, there has been an increasing interest in understanding the behavior of the human as a supervisor of complex systems [2,3,4,5,6]. More and more of the traditional control tasks are now performed by automatic systems. For example, microcomputers and minicomputers are used in a distributed data processing environment to increase the total reliability of systems[1]. Yet, the human is still an important element in most of the systems due to his unique abilities to make decisions in unexpected situations, his ability to come up with alternate solutions for problems, etc. Hence, the necessity for the human still exists, if only as a manager or supervisor. Even in completely automated systems, the human must still know how to operate such systems when failures occur unexpectedly.

* This research was supported by the National Aeronautics and Space Administration under NASA-Ames Grant NSG-2119.

In this paper, we are concerned with the human operator when he must perform different tasks simultaneously. For example, in flying an aircraft a pilot performs a variety of tasks. These tasks can be broadly subdivided into two main categories. One category is the familiar control tasks where the pilot uses the throttle, elevator, rudder and ailerons (and other control surfaces) to alter the flight conditions. Depending upon the nature of the flight, continuous attention is required of the pilot to perform the desired mission successfully. Control related actions have a direct effect on the flight conditions. Velocity, angle of attack, altitude, etc. vary continuously and are direct results of various control actions.

In contrast, there are certain other tasks which only require the pilot's attention for a certain amount of time. Included in this category are: checking the subsystems during various phases of flight, communicating with air traffic controllers, taking radio fixes, etc. These are intermittent and require a finite amount of time for successful completion. These tasks are characterized by a definite beginning and a definite ending (in time). These are called discrete tasks in this paper.

An important theoretical issue in modeling the human's behavior is to be able to predict when the discrete tasks should be performed in the time history (evolution) of the control task, i.e., this is a scheduling problem. Walden and Rouse[2] analyzed this problem with the control and discrete tasks forming a queue, to be served by the human pilot. Individual tasks were considered as customers arriving at exponentially distributed inter-arrival times. The service times had an Erlang-k distribution. Fine grained structure of the controller was not considered. Palmer[3] considered situations where the pilot must interrupt monitoring a stochastic process in order to attend to other duties, and developed normative optimal strategies for this. These strategies were compared to the results from a pilot experiment. Sheridan and Tulga[4] studied a supervisory control problem where attention was allocated dynamically between multiple processes. An optimal decision strategy was developed using dynamic programming to choose the optimal task among the possible alternatives. Muralidharan and Baron[5] are concerned with the problem of combined monitoring, decision making and control in the context of controlling multiple Remotedly Piloted Vehicles (RPVs). Human control actions are infrequent, and monitoring and decision making are the operator's main activities. A model is developed combining control and decision theoretic notions with the optimal control model. In our study, we are concerned with the problem of control, when discrete tasks must also be performed during various times. It is desired to develop a model which will explain when the discrete tasks are scheduled, and how the control task is affected by the requirement to perform discrete tasks.

THE MODEL

In an earlier paper[6] we described a heuristic model for the controller which resulted in an intermittent control activity. That model did not specifically explain when discrete tasks should be performed. It was basically a tool used to develop initial insights into the intermittent control behavior observed with the human pilots.

An analytical model is preferred to a heuristic model because, with the former, there is no need to come up with new set of heuristics for each problem. Analytical models are more general, and hence, more robust. We have employed optimal control theory to develop an analytical model. The control problem of interest is that of flying an aircraft as closely as possible to the reference trajectory. Preview of the reference trajectory is available for some distance (or equivalently, time). This is the planning horizon in which the discrete tasks must be inserted at appropriate intervals. A well-trained, well-motivated human operator is assumed to keep the errors from the reference trajectory to a minimum. He is also expected to impose a penalty on the control action, the penalty becoming higher when discrete tasks must be performed. It will be shown that the relative weights on the error and the control activity are crucial in the model.

The cost to be minimized is given by:

$$\min J = \sum_{i=1}^{N} E\left[(y_i-\tilde{y}_i)'Q_i(y_i-\tilde{y}_i) + u'_{i-1}R_{i-1}u_{i-1} \right],$$

where, $\tilde{y}_i$ - reference inputs (preview), u_i - control values, Q_i, R_i - weights.

The system dynamics is given by:

$$x_{i+1} = A_i x_i + B_i u_i, \quad y_i = C_i x_i.$$

Since preview is available, y_i are reference trajectory values into the future. An optimal control problem with preview was considered earlier by Tomizuka and Whitney[7]. In that work, the previewed information consisted of the output of a second order system driven by white noise. It was rather restrictive because no deterministic future inputs were considered. Our formulation can handle any arbitrary inputs.

The solution for this problem is obtained through dynamic programming. (See [8] for a general background.) Though a bit tedious, the solution is straight forward. The optimal control u_i is given by,

$$u_i = (B_i'K_{i+1}B_i + R_i)^{-1}B_i'[\tilde{y}_{i+1} - K_{i+1}A_i\hat{x}_i],$$

where K_i are,

$$K_N = C_N'Q_NC_N, \quad \tilde{K}_N = C_N'Q_N,$$

$$K_i = A_i' [K_{i+1} - K_{i+1}B_i(B_i'K_{i+1}B_i + R_i)^{-1}B_i'K_{i+1}] A_i + C_i'Q_iC_i,$$

$$\tilde{K}_i = A_i' [I - K_{i+1}B_i(B_i'K_{i+1}B_i + R_i)^{-1}B_i'],$$

$$\tilde{y}_N = C_N'Q_N\tilde{y}_N, \quad \tilde{y}_i = C_i'Q_i\tilde{y}_i + \tilde{K}_i\tilde{y}_{i+1}.$$

It is seen that the control u_i at time i is a function of the feedback of the state (or more appropriately, its estimate) at that time, and a feedforward of the reference trajectory.

The important feature in scheduling control and discrete tasks is the possibility of directly influencing the control through the weights R on control (see equation for u_i above). When it is desired to insert a discrete task somewhere in the planning horizon, the control weighting is increased in that interval, resulting in a reduced control value. If this is done for the entire planning horizon (i.e., altering the weights in intervals of desired discrete task activity), this has the desirable effect of redistributing control activity to keep the overall cost at an optimal value.

To schedule discrete tasks, the control time history is first calculated for the entire planning horizon with constant weights on control. Since states (x_i) are not known, they are estimated using the calculated controls at the previous time instant. The number of discrete tasks to be performed in the planning horizon are determined, and the time required for performance of each is estimated. Relative minima of control activity are chosen for each discrete task to be inserted, and the weights on control are increased over these intervals. Then, a new control history is calculated over the planning horizon. Since weights or penalties have been increased in these intervals, control activity is depressed. The control calculated thus is optimal for the combined control-discrete task problem.

From this model, it is possible to insert discrete tasks when the need arises. Since one model gives the control behavior as well as the time intervals when discrete tasks are to be performed, fractions of time spent on each type of task can be calculated if the nature of the tasks is known.

EXPERIMENT

An experiment is designed to test the validity/appropriateness of this model in a multitask environment.* A VT11 graphics display driven by a PDP 11/40 minicomputer interfaced with an analog computer (Applied Dynamics, AD5) is used. The control task involves flying over a map at a constant altitude and airspeed using aileron control to turn. (See Figure 1). The proper amount of rudder is automatically provided by the computer for a coordinated turn. A track up display is used where the map to be flown moves past a fixed airplane symbol. The reference trajectory for a few minutes into the future is available for preview. A modified form of an artificial horizon with a moving wing over a fixed horizon is used to indicate the attitude. Aileron angle is controlled using a joystick.

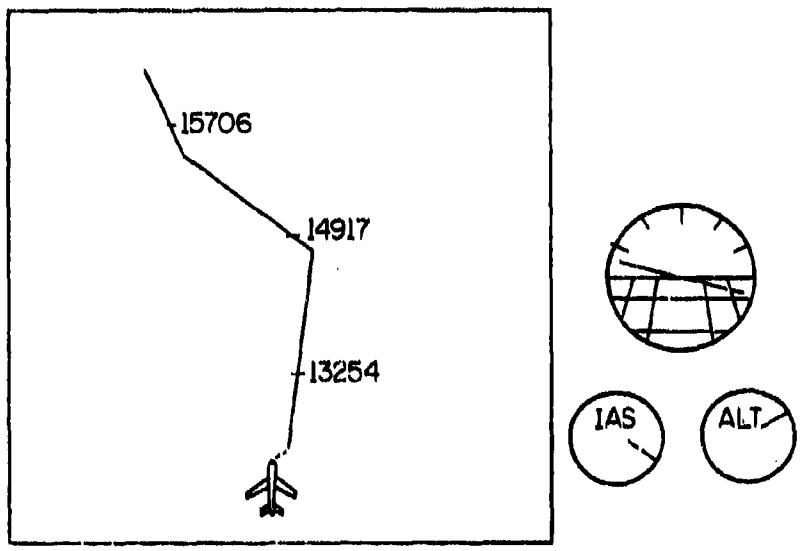


Figure 1

Discrete tasks are presented as number entry tasks. At various times, while flying along the map, five digit numbers appear and move along the map. Successful performance of these discrete tasks requires correct entry of the numbers before they disappear. A key pad is used to enter the numbers. Some or all of these digits can be entered at one time. A switch near the

* The authors wish to thank Everett Palmer of NASA Ames Research Center for his constructive suggestions in the development of the experiment.

joystick is used to initiate/terminate the number entry tasks. When a decision is made to enter a number, this switch is pressed, and the map disappears from the screen. Entry of each digit results in a 'beep'. After correctly entering a number, the next number cannot be entered immmediately. The subject must wait until the small horizontal tic mark corresponding to the number just entered goes past the airplane symbol.

The subjects are expected to 'fly' the aircraft such that the airplane remains on the map as closely as possible. They are instructed to schedule the number entry tasks, so that both tasks can be performed acceptably and comfortably, without either task requiring too much effort. If an error occurs in entering the number, even in one digit, the entire number must be re-entered. The map remains blanked out during data entry. Even when the map is not seen, however, the airplane does not stop. Thus, subjects try to perform the discrete tasks quickly.

Subjects are trained in the control task until their performance reaches a consistent value. When there is no further improvement in the mean-squared error, discrete tasks are performed along with the control task. Along with some graduate students in engineering, some private pilots are being used as subjects. Discrete tasks are scheduled at constant arrival rates, as well as with a Poisson distribution with the mean equal to that of the constant arrival rate. Experiments are in progress at this time.

REFERENCES

1. "Electronics Can Help Prevent More Nuclear Accidents", Editorial in Electronics, April 12, 1979, p24.

2. Walden, R. S.; and Rouse, W. B.: "A Queueing Model of Pilot Decisionmaking in a Multitask Flight Management Situation", IEEE Trans. Syst., Man, Cybern., vol. SMC-8, no. 12, pp. 867-875, December 1978.

3. Palmer, E.: "Interrupted Monitoring of a Stochastic Process", Proceedings of the Thirteenth Annual Conference on Manual Control, MIT, June 15-17, 1977.

4. Sheridan, T. B.; and Tulga, M. K.: "A Model for Dynamic Allocation of Human Attention Among Multiple Tasks", Proceedings of the Fourteenth Annual Conference on Manual Control, University of Southern California, April 25-27,1978.

5. Muralidharan, R.; and Baron, S.: "Combined Monitoring, Decision and Control Model for the Human Operator in a Command and Control Task", Proceedings of the Fourteenth Annual Conference on Manual Control, University of Southern California, April 25-27, 1978.

6. Govindaraj, T.; and Rouse, W. B.: "Modeling the Human as a Controller in a Multitask Environment", Proceedings of the Fourteenth Annual Conference on Manual Control, University of Southern California, April 25-27.

7. Tomizuka, M; and Whitney, D. E.: "The Human Operator in Manual Preview Tracking (an Experiment and Its Modeling Via Optimal Control)", Transactions of the ASME, Journal of Dynamic Systems, Measurement, and Control, December 1976, pp. 407-413.

8. Bertsekas, D. P.: "DYNAMIC PROGRAMMING AND STOCHASTIC CONTROL", Academic Press, 1976.

A STUDY OF HUMAN ATTENTION ALLOCATION IN MULTIPLE PROCESS MONITORING SITUATIONS*

Joel S. Greenstein and William B. Rouse

Department of Mechanical and Industrial Engineering
Coordinated Science Laboratory
University of Illinois
Urbana, Illinois 61801

ABSTRACT

The human operator's allocation of attention among multiple dynamic processes is considered. An experimental investigation of the human's attention allocation performance in a specific multiple process monitoring situation is discussed. In this experiment, subjects simultaneously monitored nine dynamic processes for the occurrence of failures and made decisions regarding the processes to which they wished to allocate attention. Parameters associated with each of the processes were the mean time between failures of the process, the mean time required to repair the process, and the cost per unit time of delay in servicing the process. The effects of these parameters on the timing and ordering of the subjects' attention allocation decisions are investigated and an approach to the modeling of the subjects' attention allocation performance consistent with these effects is presented. The goal of this work is the development of a model of multi-task decision making which could be used in the design and implementation of computer-aided decision making systems.

INTRODUCTION

In many systems, the human operator spends much of his time monitoring subsystems for events which call for action on his part. As the complexity of these systems increases, the operator becomes responsible for more subsystems of greater variety. There is consequently a greater probability that the operator will encounter situations in which there are more tasks than he can acceptably perform.

One means of maintaining the operator's workload at a satisfactory level is the introduction of automation capable of performing some of the operator's tasks. Models of the

*This research was supported by the National Aeronautics and Space Administration under NASA-Ames Grant NSG-2119.

operator's task performance would be of use in predicting the performance gains to be expected from the introduction of such aids. Further, in systems in which the responsibilities for some tasks are shared by the operator and an automated decision maker, these models might also be used within the system to coordinate the actions of the two decision makers.

Senders [1] and Smallwood [2] have modeled human decision making in multiple process monitoring tasks. Senders postulated that the human monitor samples his displays in a manner which allows reconstruction of the displayed signals. An information theory approach is employed to determine how often and for what duration the human must sample each display. Smallwood proposed that the human operator forms an internal model of the processes he is monitoring and of the environment relevant to his task as a result of his past perceptions of them. A situation is considered in which the operator seeks to detect excursions of instruments beyond threshold values. The operator is modeled as directing his attention to the instrument whose current probability of exceeding threshold (based on the operator's internal model) is greatest.

Carbonell [3,4] and Senders and Posner [5] have proposed queueing theory approaches to the modeling of human decision making in multiple process monitoring tasks. Carbonell uses a priority queueing discipline. He assumes that the human operator attempts to minimize the risk involved in not observing other instruments when he chooses to monitor a particular instrument. Senders and Posner employ a first-come first-served service discipline. They suggest two models which might be used to estimate the inter-observation intervals for an instrument (i.e. the time between arrivals of the instrument to the queue of instruments awaiting observation by the human monitor). The first model involves the degree of the observer's uncertainty about the value of the variable displayed on the instrument. The second model involves the probability that the displayed variable will exceed an acceptable limit.

The models cited above emphasize the monitoring of displays, rather than the decisions or actions that result from the human operator's perception of the displayed values. The operator's motivation for monitoring the displays is the possibility that an event which requires his action will occur. The multi-task decision making problem addressed in this paper concerns the event detection and action selection decisions the operator makes on the basis of the information he gains through monitoring.

Human decision making in such multi-task situations, then, might be modeled in terms of the manner in which the human detects events related to his tasks and the manner in which he allocates his attention among his tasks once he feels events have occurred. Gai and Curry [6] have developed a model of the human monitor in a failure detection task. The model has two stages, the first a Kalman filter which estimates the states and observations of the monitored process, and the second a decision

mechanism which operates on the Kalman filter residuals using sequential analysis concepts. The model can be used to describe the human monitor's detection of additive failures in stationary random processes. Sheridan and Tulga [7] have modeled the manner in which the human operator allocates his attention among various tasks. They address a situation in which events present themselves unequivocally and use a dynamic programming approach to determine the action sequence which maximizes the operator's earnings. Execution of this action sequence is begun, but can be superceded by execution of a new sequence calculated in response to the appearance of additional tasks.

Rouse [8] has investigated the issue of allocation of decision making responsibility between a human operator and an automated decision maker. He presents a mathematical formulation of multi-task decision making situations appropriate to modeling either decision maker. Based on displayed information, the decision maker is assumed to generate probabilities that events have occurred in his tasks. He also generates density functions which characterize his perceptions of what might occur in his tasks while his attention is diverted to a particular task and how long his attention will be diverted should he decide to take a given action. Given the estimates of event occurrence probabilities and the perceptions of density functions of interarrival and action times, the decision maker is assumed to choose actions so as to minimize an appropriate cost criterion.

This formulation has been applied to the design of a computer-aided decision making system within a queueing theory framework [9,10]. Two experiments within a flight management context were used to illustrate the value of this multi-task formulation. It was found that the queueing theory model provided a quite reasonable description of multi-task decision making in terms of task waiting times and utilization (i.e., fraction of time busy) of decision makers. The nature of the flight management task employed in this research, however, allowed avoidance of the issue of how the human decision maker detects events.

Greenstein and Rouse [11] have proposed a model of the human's event detection performance in a multi-task situation and have applied this model to a specific multiple process monitoring situation. An assumption of the model is that, in attempting to detect events, the human generates estimates of the probabilities that events have occurred. Discriminant analysis is used to model the human's generation of these probability estimates. An experimental study of human event detection in a multiple process monitoring situation was done. This study employed a situation in which subjects simultaneously monitored several dynamic processes for the occurrence of events and made yes/no decisions on the presence of events in each process. The event detection model was shown to provide a good fit to the data obtained in the experimental study.

In this experimental study of event detection, subjects were re,uired to make yes/no decisions on the presence of events in the processes they monitored. In many multi-task situations the human operator is not forced to make such yes/no decisions with respect to each of his tasks. Instead, he uses his estimates of the probabilities of task-related events in deciding how to allocate his attention among tasks. This paper describes an experimental study of the human's attention allocation performance in a multiple process monitoring situation similar to the one employed in our investigation of event detection.

AN ATTENTION ALLOCATION EXPERIMENT

Fig. 1 illustrates the display observed by subjects in the experiment. This static display was generated on a Tektronix 4010 by a time-shared DEC-System 10 and depicts the measured values of the outputs of nine processes over 100 sampling intervals (i.e. 101 points). The processes had identical second order system dynamics with a natural frequency of 0.75 rad/sec and a damping ratio of 0.5. Samples were taken at 0.2 second intervals. The inputs to the processes were zero-mean Gaussian white noise sequences of identical variance. The displayed measurements were obtained by corrupting the process outputs with additive zero-mean Gaussian white noise sequences which normally had identical variance. The measurement noise variance was normally selected to yield measurements with signal-to-noise

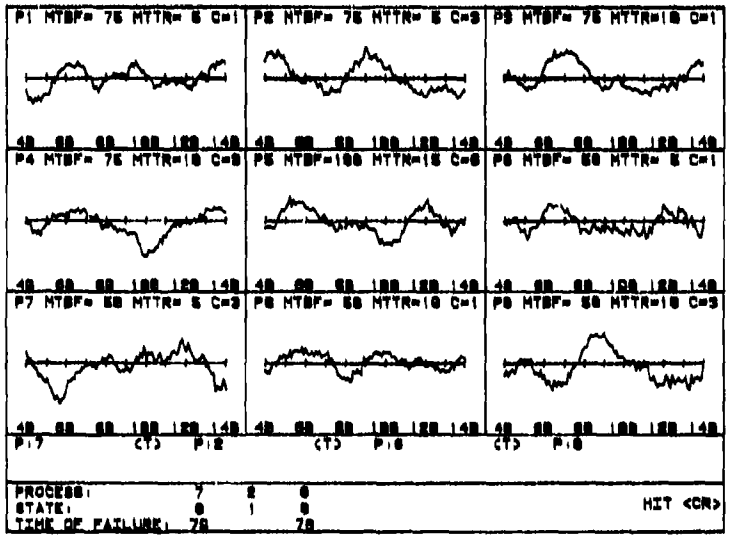


Fig. 1. The multiple process monitoring situation. (Processes 1, 5, 6, 7, and 9 have failed at times 131, 95, 76, 79, and 91, respectively. Processes 7, 2, and 6 have been selected for service.)

ratios of 25.0. A process failure was defined by a gradual increase in the measurement noise variance such that the signal-to-noise ratio of each measurement was decreased to 97.5% of the signal-to-noise ratio of the immediately preceding measurement. A process failure manifests itself, then, by an exponential decay of the signal-to-noise ratio for measurements following the time of failure. Thus, failures become more pronounced with each measurement following their occurrence. This is illustrated in Fig. 2 in which all nine processes have failed at various points within the time range depicted on this display. The points at which these failures occurred have been denoted in the figure by solid vertical lines.

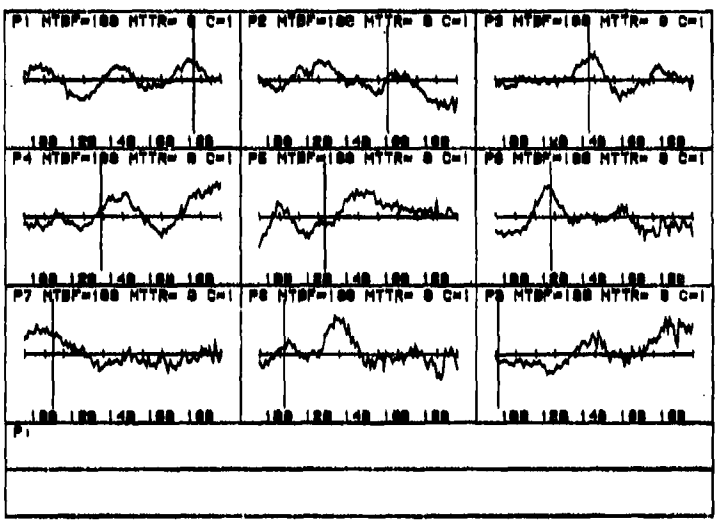


Fig. 2. A display in which all processes have failed. (The failure times are denoted by solid vertical lines. These lines did not appear on the display during the experiment.)

The task of the subject in this experimental situation was to service, or allocate attention to, processes in a manner which minimized the total cost incurred due to delays in the servicing of failed processes. Some of the processes tended to fail more frequently than others. Servicing a process implied a diversion of attention from the other processes and the amount of time required to service a process varied across the processes. The cost of delaying service to a process also differed from process to process. These aspects of the situation were quantified by associating a set of three parameters with each of the nine processes -- the mean time between failures of the process (MTBF), the mean time to repair the process (MTTR), and the cost per unit time of delay in servicing a failed process (C). The values of these parameters for each process were included in the display presented to the subject.

After scanning the nine process histories depicted on the display, the subject was given an opportunity to key in the numbers of processes to which he had decided to allocate attention. He also used a graphic cursor to enter estimates of the times at which he believed failures had occurred. He completed his entries by entering a zero. The processes he entered were then serviced in the order in which he had entered them. The first process entered by the subject was serviced over an interval which began with the last time point displayed on the screen and extended forward the number of time units equal to the MTTR associated with the process. At the completion of this service, the process was reset to normal if it had failed before this point and service was begun on the next process entered by the subject. When all processes entered had been serviced, the subject was given feedback regarding the state of each serviced process at the instant before servicing of that process was completed ("1" indicating the normal state, "0" indicating the failed state). If a serviced process had in fact been in the failed state, the time of failure was also indicated. An iteration in the trial ended with erasure of the display and scoring of the subject's performance on the iteration.

Another iteration was then begun by generating a new display depicting the process histories advanced in time by an amount equal to the sum of the service times of the processes entered on the last iteration plus an additional increment of 5 time units. The display which would follow that shown in Fig. 1 is illustrated by Fig. 3.

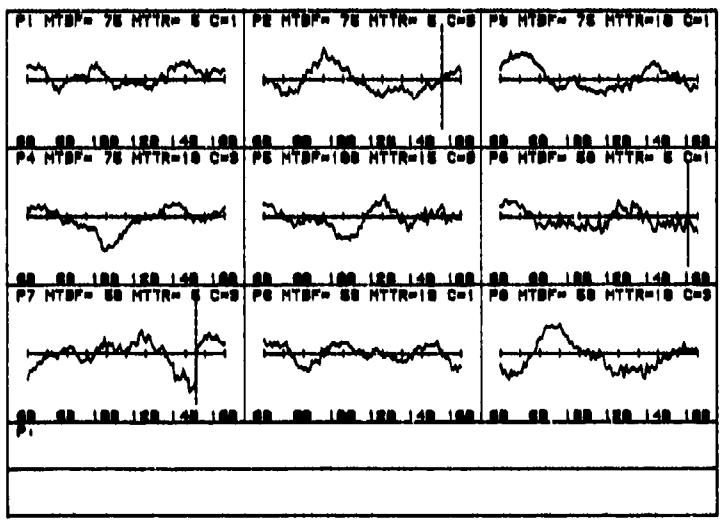


Fig. 3. The display which would follow that shown in Fig. 1. (The dashed vertical lines indicate to the subject the point at which service of each process was last completed.)

Two subjects were given six trials with this experimental situation over a period of three to four days. Each trial consisted of 25 iterations (an initial display followed by 24 updated displays) and was of approximately 45 minutes duration. The parameter values associated with each of the nine processes were as depicted in Figures 1 and 3. These values were held constant across all trials, but the actual interarrival times used to schedule failures on the successive trials given a subject differed from trial to trial. The interarrival times of the failures scheduled to occur in a process were exponentially distributed with a mean equal to to the MTBF associated with that process. Not all scheduled failures actually occurred. If a failure was scheduled to occur in a process in which a previous failure had not yet been attended to, the scheduled failure was deleted from the trial.

Prior to beginning the six trials with the experimental situation described above, the subjects were given five trials with a situation in which the decision making tasks were somewhat simpler. In these trials MTTR was set to zero and C to one for all processes. Each updated display depicted the process histories advanced 10 time units from the preceding display. With zero MTTR and equal C for all processes, the decision to allocate attention to one process involved no diversion of attention from the other processes and a delay in servicing one process was no more costly than a delay in servicing any other process. After scanning the process histories depicted on the display, the subject simply entered the numbers of processes in which he had decided failures had occurred. These trials provided the subjects with experience in the detection of process failures without the additional task of deciding how to allocate attention among processes with different attentional requirements and costs. Subjects were also given several iterations of training before these trials during which solid vertical lines were included on the process histories to mark exactly when and where events had occurred (see Fig. 2).

RESULTS

Two analyses of variance were performed on the attention allocation decisions made by each of the two subjects on the last four trials of the experiment. The first of these analyses examined the effects of MTBF, MTTR, and C on the number of time units which elapsed from the time at which a process failed until the time at which the subject responded to the failure. The second analysis examined the effects of the three parameters on the number of time units which elapsed from the time at which the subject keyed in a process number to the time at which service of the process was initiated. The analyses did not include the attention allocation decisions made with respect to process 5. The parameter values associated with the other eight processes comprised a factorial design with two levels for each of three

factors (MTBF, MTTR, and C).

Table 1 presents the results of the analysis of variance for the elapsed time from failure of a process until response to that process. The analysis indicates that, of the three parameters studied, only the costs associated with the processes had a significant effect on the time at which the subject responded to the failure of a process. This was true for both subjects (S1 and S2). The analysis also indicates that none of the interaction effects of the three parameters were statistically significant.

Table 1. Time until response to a failure

Analysis Of Variance				
Source Of Variation	Degrees Of Freedom		F_0	
	S1	S2	S1	S2
MTBF (A)	1	1	0.66	2.14
MTTR (B)	1	1	0.	0.02
Cost (C)	1	1	5.28 *	6.00 *
AB	1	1	0.25	0.64
AC	1	1	0.18	2.14
BC	1	1	0.24	2.02
ABC	1	1	3.60	0.01
Error	96	128		
* Significant at the 2.5% level				

Table 2 presents the results of the analysis of variance for the elapsed time from the subject's entry of a process number to the initiation of service of the process. This analysis indicates that the costs associated with processes selected for service on a given iteration of the experiment had a significant effect on the order in which these processes were scheduled for service. This was true for both subjects. The analysis also indicates that for S1 the MTTRs associated with the processes also had a significant effect on the subject's ordering of processes for service on a given iteration. Again, none of the interaction effects of the parameters were statistically significant.

Table 2. Time from process entry to initiation of service

Analysis Of Variance				
Source Of Variation	Degrees Of Freedom		F_0	
	S1	S2	S1	S2
MTBF (A)	1	1	0.40	0.15
MTTR (B)	1	1	17.59 *	1.36
Cost (C)	1	1	14.25 *	29.63 *
AB	1	1	0.04	0.15
AC	1	1	1.10	0.60
BC	1	1	0.70	0.27
ABC	1	1	0.40	0.27
Error	104	168		

* Significant at the 1% level

A MODELING APPROACH

Analysis of the attention allocation performance in the specific multiple process monitoring situation presented in this paper indicated that the decision to attend to a process was influenced by the cost associated with the delay of service to the process. The mean time between failures in the process and the mean time required to repair the process did not appear to exert a strong influence on this decision. Once the decision had been made to allocate attention to several processes, the order in which these processes were attended appeared to be influenced by both the cost of delay of service and the mean time to repair associated with the process. The mean time between failures in the process again appeared not to exert a strong influence.

The attention allocation performance of a human operator in this situation might then perhaps be modeled using a two stage procedure. In the first stage, the operator estimates the probability that each of the processes has failed based on his observation of the displayed process histories. If this probability exceeds a threshold, the operator includes this process among those for which service must be scheduled. The value of this threshold would depend on the costs associated with the processes.

In the second stage of the attention allocation procedure, the operator must determine the order in which the processes selected in stage 1 are to be serviced. One possible approach might be to service these processes in order of highest to lowest C/MTTR. That is, of the processes selected for attention in

stage 1, that process for which the associated cost of delay divided by the mean time to repair is highest would be serviced first. The process for which this quotient is lowest would be serviced last. This approach is motivated by the fact that the experimental situation can be modeled in queueing terms as an (M/G/1) queue with a nonpremptive priority service discipline and finite queue capacity. Failures occur in any of nine processes and await service by the operator. Cox and Smith [12] have shown that for an (M/G/1) queue with a nonpreemptive priority service discipline, the servicing of customers in order of highest to lowest C/MTTR is optimal in the sense of minimizing the expected cost due to delays in service. This simple result does not hold, however, for a system with finite queue capacity. Mova and Ponomarenko [13] have shown that the optimal priority assignment for such a system varies with the instantaneous configuration of the queue.

The queueing results cited also assume that customers present themselves unequivocally. That is, the probability that a customer is present is either zero or one. In the process monitoring situation considered in this paper, failures do not present themselves unequivocally. The operator can seldom say with absolute certainty that a process has or has not failed. The model of event detection performance proposed by Greenstein and Rouse [11] has been applied to a multiple process monitoring situation very similar to the one described in this paper and can be used to generate estimates of the probabilities of process failures. The two stage model of attention allocation performance suggested here allows those processes for which the failure probability exceeds threshold in stage 1 to be treated in stage 2 as customers which have presented themselves unequivocally.

CONCLUSION

The human operator's allocation of attention among multiple dynamic processes has been considered. An experimental investigation of the human's attention allocation performance in a specific multiple process monitoring situation was described. In this experiment, subjects simultaneously monitored nine dynamic processes for the occurrence of failures and made decisions regarding the processes to which they wished to allocate attention. Parameters associated with each of the processes were the mean time between failures of the process, the mean time required to repair the process, and the cost per unit time of delay in servicing the process. Analysis of the subjects' attention allocation performance in this experimental situation indicated that the decision to attend to a process was influenced by the cost associated with the delay of service to the process. The mean time between failures in the process and the mean time required to repair the process did not appear to exert a strong influence on the decision. Once the decision was

made to allocate attention to several processes, the order in which these processes were attended appeared to be influenced by both the cost of delay of service and the mean time to repair associated with the process. The mean time between failures in the process again appeared not to exert a strong influence.

A two stage procedure was suggested as an approach to modeling the operator's attention allocation performance in this situation. A set of processes to which attention is to be allocated is selected in the first stage. Selection of a process for attention is based on an estimate of the probability that the process has failed and on the costs associated with the processes. The order in which the processes selected in stage 1 are serviced is determined in the second stage of the procedure. This order might be chosen to minimize the expected cost due to delays in service of the processes selected in stage 1. It is also possible that a simpler ordering strategy achieving low, but not minimal expected costs might be employed.

We now plan to continue our experimental investigation of attention allocation performance in the multiple process monitoring situation described in this paper. Of the three parameters manipulated in the experiment discussed here, the costs and the mean times to repair associated with the processes appear to be most interesting in terms of their effects on the operator's attention allocation decisions. We are therefore planning an experiment in which three levels of each of these two factors are employed across the nine processes of the display. The results of this investigation will be used to refine and validate a model of human attention allocation which could be used in the design and implementation of a computer-aided process monitoring system.

REFERENCES

1. Senders, J. W., "The Human Operator as a Monitor and Controller of Multidegree of Freedom Systems," IEEE Trans. Human Factors in Electronics, vol. HFE-5, no.1, pp.2-5, Sept. 1964.

2. Smallwood, R. D., "Internal Models and the Human Instrument Monitor," IEEE Trans. Human Factors in Electronics, vol. HFE-8, no.3, pp. 181-187, Sept. 1967.

3. Carbonell, J. R., "A Queueing Model of Many-Instrument Visual Sampling," IEEE Trans. Human Factors in Electronics, vol. HFE-7, no.4, pp. 157-164, Dec. 1966.

4. Carbonell, J. R., Ward, J. L., and Senders, J. W., "A Queueing Model of Visual Sampling: Experimental Validation," IEEE Trans. Man-Machine Systems, vol. MMS-9, no.3, pp.82-87, Sept. 1968.

5. Senders, J. W. and Posner, M. J. M., "A Queueing Model of Monitoring and Supervisory Behavior," in T. B. Sheridan and G. Johannsen, Eds., Monitoring Behavior and Supervisory Control, New York: Plenum 1976, pp.245-259.

6. Gai, E. G. and Curry, R. E., " A Model of the Human Observer in Failure Detection Tasks," IEEE Trans. Systems, Man, and Cybernetics, vol. SMC-6, no.2, pp.85-94, Feb. 1976.

7. Sheridan, T. B. and Tulga, M. K., "A Model for Dynamic Allocation of Human Attention Among Multiple Tasks," Proceedings of the Fourteenth Annual Conference on Manual Control, University of Southern California, April 1978, NASA CP-2060, pp. 569-592.

8. Rouse, W. B., "Human-Computer Interaction in Multitask Situations," IEEE Trans. Systems, Man, and Cybernetics, vol. SMC-7, no.5, pp. 384-392, May 1977.

9. Walden, R. S. and Rouse, W. B., "A Queueing Model of Pilot Decision Making in a Multi-Task Flight Management Situation," IEEE Trans. Systems, Man, and Cybernetics, vol. SMC-8, no. 12, Dec. 1978.

10. Chu, Y. and Rouse, W. B., "Pilot Decision Making in a Computer-Aided Flight Management Situation," Proceedings of the Fourteenth Annual Conference on Manual Control, University of Southern California, April 1978, NASA CP-2060, pp. 677-690.

11. Greenstein, J. S. and Rouse, W. B., "A Model of Human Event Detection in Multiple Process Monitoring Situations," Proceedings of the Fourteenth Annual Conference on Manual Control, University of Southern California, April, 1978, NASA CP-2060, pp. 663-675.

12. Cox, D. R., and Smith, W. L., Queues, London: Methuen, 1961.

13. Mova, V. V. and Ponomarenko, L. A., "On the Optimal Assignment of Priorities Depending on the State of a Servicing System with a Finite Number of Waiting Places," Engineering Cybernetics, Vol.12, No.5, Sept.-Oct. 1974, pp. 66-72.

MODELS FOR HUMAN OPERATORS BASED ON NONUNIFORMLY SAMPLED SYSTEMS

George D. Kontopidis
Filson H. Glanz
David E. Limbert

Kingsbury Hall
University of New Hampshire
Durham N.H. 03824

Abstract

Properties of the human operator and results of previous studies indicate reasons to use nonuniformly sampled systems for modeling man-machine systems.

Modern estimation and identification theory is used to establish a framework for the analysis of nonuniformly sampled systems. The irregularly sampled version of a time invariant, linear, stochastic, dynamic system is used for quantitative study. No assumptions on the sampling sequence are imposed but sampling intervals have to be known explicitly. The human operator is partially modeled by a state estimator which uses nonuniform samples of the plant output. When the parameters of the supervised system are partially known, the irregularly spaced observations of the input and output signals are used for both parameter identification and state estimation. Sequential and off-line algorithms have been developed for the identification and estimation processes.

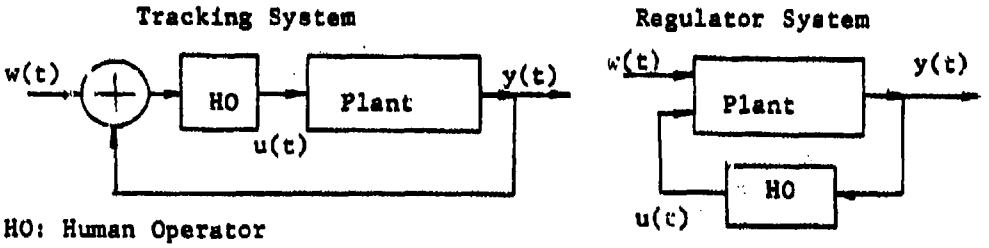
In this paper we focus on theoretical development of state estimation and parameter identification methods for nonuniformly sampled systems which can be applied in modeling of manually controlled systems.

I. Introduction

The basic goal in developing models of the human operator is to systematically and easily predict pilot behavior and system performance in complex control tasks. After the model has been derived, properties of the composite man-machine system can be investigated. Researchers' interest usually focuses on the following:

- To determine the overall dynamics of the man-machine systems.
- To specify the man to machine task.
- To improve the overall system performance by allocating the operator's attention and actions.

In this introductory part we summarize the major studies in manual control systems (that is, systems where a human operator is a part of the closed loop system according to Fig. 1).



HO: Human Operator

y(t): Displayed output w(t): Noise & Disturbances

u(t): Manual Controls

Fig. 1: Manually Controlled Systems

This selected summary is then used to explain the reasons for using nonuniformly sampled system models for the operator. In the next sections (II, III, and IV) we present the results of a theoretical study on nonuniformly sampled systems.

The main approaches in modeling the human operator have been the frequency response describing function as applied by Mc Ruer and Weir [25] and the optimal control approach used by Kleinman, Baron and Levison [16]. The latter approach was applied to the hover control of a VTOL-type vehicle [16]. The same model was applied to the multiaxis manual control of a fighter aircraft by Harrington [13]. Repa and Wierwille [32] presented experimental results which were aimed at determining the influence of the vehicle control characteristics on driving performance and comfort. The helmsman of a large ship was modeled and simulated successfully by Veldhuyzen [41]; in his approach frequency domain ideas were used to optimize the model parameters. A very interesting study in modeling the human operator as a supervisor of a distributed computer network was done by Tuga and Sheridan [40]; this is a typical example of modeling the hierarchial control of large scale systems including a human as a decision maker. Rouse's [33] dissertation and his later studies [34] in modeling the human by using modern estimation theory approaches and hypothesis testing, is a good theoretical background for improving existing models. The experiments by Enstrom [8] to find the attention allocation between control and monitoring tasks are also based on hypothesis testing; to assume more than one model and choose the one that best

fits the experimental data.

Application of identification methods by Gabay and Merhav [10] to a closed loop task gives a closed form result for the transfer function of the human operator; the result is

$$G(s)=5.5\exp(-0.19s)\ \frac{0.25s+1}{0.35s+1}\ .$$

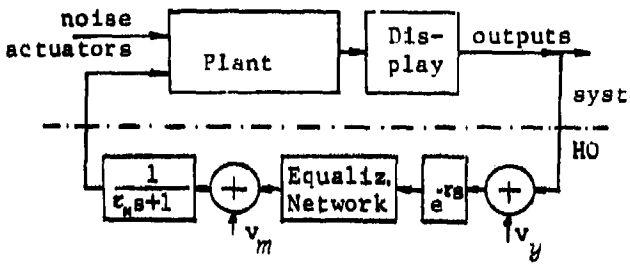
Such results depend upon the particular experiment; for example, Wijk and Kok [44] suppress the exponential term when the plant is relatively slow. Recent studies in the area of improving models of humans by using the position of the eye were presented by Repperger and Hartzell [31]; the effect of work load on the variances of the observation and motor noise [16] has been studied partially in [42] and [43].

In the following, we introduce the assertion that the human operator can be modeled as a state estimator of a nonuniformly sampled system. Other researchers [34], [7] have demonstrated, with some success, that the operator does construct an internal model of the phenomena with which he is concerned (a Kalman filter), so he can predict its behavior and control it. The basic idea behind the assertion of using nonuniformly sampled systems for such models is that the operator does not observe the output of the system "continuously" but "samples" the output at discrete time instants; at the same time, he manually controls the plant in a discrete manner.

The first indication of the desirability of considering nonuniformly sampled models for the plant comes from the Kleinman et al. [16] model. The human operator controls a linear, time invariant, stochastic system optimally, subject to his inherent limitations. These limitations are considered to be:

- A time delay (τ), representing cognitive, visual central processing and neuromotor delays.
- "Remnant" signals, divided into an observation noise ($\{v_y(t)\}$ with covariance V_y) to represent signal degradation due to work load, scanning effects and signal thresholds, and a motor noise ($\{v_m(t)\}$ with covariance V_m) to represent random errors in executing the intended control.
- A "neuromuscular lag" (τ_N) to represent neuromuscular dynamics.

A block diagram of their model is illustrated in Fig. 2. In this model the estimator is designed by using a continuous time Kalman filter or a discrete time invariant one. If we assume that the visual sampling rate (of the displayed outputs) is not constant, the estimator needs modifications. There are several



v_m: Motor Noise with covariance V_m.

v_y: Observation Noise with covariance V_y.

HO: Human Operator.

Fig. 2: Human Operator Model [16]

studies on such visual sampling processes [47] and experiments on the sensitivity of the eye that can be used to adapt the original model. As an example, Alen and Ruer [2] have found that a driver's delay (τ) increases at lower vehicle speeds due to decreasing speed. Therefore, if the irregularity of the visual sampling rate is taken into account, the model can be improved.

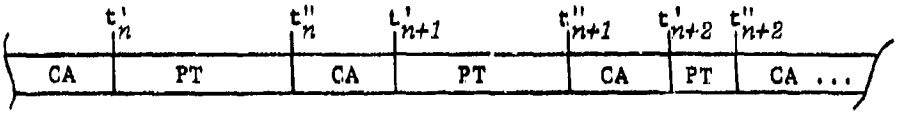
The second indication of the irregularity of the sampling comes from Rouse [34], Kok [17] and Kvalseth [15]. They underline the hypothesis that the human supervisor acts partially as observer/reconstructor and partially as controller/decision maker. The following hypothesis are also used:

- The model of the observer depends on the supervised system, the display and the noise parameters.
- The controller part depends only on the supervised system and can be considered independent of the observer part.

Now, if we impose the hypothesis that

- the human supervisor acts as a discrete process with a unique processing unit,

the observation rate depends on the time it takes to process a given piece of information and the time it takes to realize the controlling action. This fact is illustrated in Fig. 3.



CA: Control Action Time

PT: Information Processing Time

t'_n: Observation Sampling

t''_n: Controlling Sequence

Fig. 3: Process and Control Sequence Timing

Another indication of sampling irregularity comes from systems containing computers as well as humans. In large scale systems (e.g. computer networks) the time between the lower priority tasks is usually irregular because of the asynchronous interrupt configuration. For modeling such systems, a nonuniformly sampled model is appropriate. To be more specific, we refer to the task of a pilot controlling an aircraft by giving appropriate commands to a computer (Palmer [27]). In this supervisory role, the pilot must allocate his attention between monitoring the aircraft performance and giving commands to the computer (see also the simulation results by Enstrom and Rouse [8]). In this example an interruption in observing the flight measurements due to attention to "other duties" or "side tasks", causes nonuniform sampling of observations.

The above reasons explain why the controlled plant is actually observed and controlled nonuniformly. Thereafter the following questions arise:

- How can we find the sampling instants $\{t_0, t_1, \ldots\}$; in other words, what are the statistics of the intervals $T_n = t_{n+1} - t_n$?
- What modifications are necessary to existing models if we assume nonuniform sampling?

In this paper we deal with a part of the above subject. The class of systems and the assumptions for the following presentation are:

- The controlled plant is a multiple input-output, linear, time invariant observable and controllable. The parameters of the plant are not necessarily known "exactly"; this means that parameter identification also takes place.
- Additive noise with known statistics corrupts the measurements, the controlled plant, and the control input.
- Both the control inputs and the observed outputs are sampled simultaneously at $t_n; n \in N_0$* and the intervals $T_n = t_{n+1} - t_n; n \in N_0$ are assumed to be known explicitly. There are no assumptions on the length of T_n.

In reality, the observation sequence $\{t_n'\}$ can be found by sensing the position of the eye with respect to the display and the controlling sequense $\{t_n''\}$ can be found by sensing the changes of the actuators. We realize that the above schemes are not necessarily the best ones (e.g. the human might be observing the display and performing his control task as in [8]). Generalization to a broader class of systems and relaxation of the last restrictions ("simultaneously" and "interval known explicitly") are subjects of future work.

* The set N_0 is defined by: $N_0 = \{0, 1, \ldots\}$. Similarly, $N = \{1, 2, \ldots\}$,

The results of the analysis of nonuniformly sampled systems are presented in the following manner. In Sec. II the linear system is sampled nonuniformly and modeled. In Sec. III we refer to state estimators for the nonuniformly sampled systems; at the end of the paper in the appendix, computational aspects related to the real time implementation of the estimators are presented. In Sec. IV we assume that the original linear system is partially known; then we try to estimate the states of the system and to identify the parameters using irregularly taken samples. Finally in Sec. V we indicate areas of future work and possible extentions. Most of the proofs have been omitted but the interested reader can find them in [18].

II. Models and Simulation Algorithms

In this section the nonuniform sampling process is studied. A linear, time invariant, stochastic model is sampled at time instants $t_n : n \in N$. The objective here is to find the difference equations which relate input and output samples.

The state space representation of the physical system (plant) is assumed to be the following stochastic differential equation:

$$\begin{aligned} dx &= F_s x dt + G_{1s} u dt + G_{2s} dw \\ dy &= C_s x dt + D_s u dt + dv \end{aligned} \tag{S}$$

where F_s is an nxn real matrix, G_{1s} and G_{2s} are $nx\ell_1$ and $nx\ell_2$ matrices respectively, C_s is an mxn observation matrix, and D_s is given by CG_{1s}. The n-, m-dimensional stochastic processes $\{x(t)\}$ and $\{y(t)\}$ are the states and the observations (outputs) respectively. The plant noise and the measurement noise are modeled by ℓ_2- and m-dimensional Wiener-Levy uncorrelated processes with incremental covariances $E\{dwdw^T\}=Qdt$ and $E\{dvdv^T\}=Rdt$ respectively. The control input u(t) is assumed to be an ℓ_1-dimensional deterministic signal. The above system S comes from the superposition of a deterministic and a stochastic system (S_1 and S_2 respectively), and is specified by the equations:

$$\begin{aligned} \dot{x}_1 &= Fx_1 + G_1 u \\ y_1 &= Cx_1 \end{aligned} \tag{S_1}$$

and

$$\begin{aligned} dx_2 &= F\, x_2 dt + G_2 dw \\ dy_2 &= C_3 x_2 dt + dv \end{aligned} \tag{S_2}$$

with $F_s=F,\quad G_{1s}=G_1,\quad C_s=CF,\quad D_s=CG_1,\quad G_{2s}=G_2$

and $dx=dx_1+dx_2$

$dy=dy_1+dy_2$

The nonuniform sampling process is studied separately for the deterministic system S_1 and the stochastic system S_2. The sampled version of S can be found by solving the state equations from t_n to t_{n+1}. The result is a time varying discrete model SD given by:

$$x_1(t_{n+1})=A(n)x_1(t_n)+B(n)u(h_n) \qquad \text{(SD)}$$

$$y_1(t_n)=Cx_1(t_n)$$

where h_n is a particular point between t_n and t_{n+1}, and the matrices A(n) and B(n) are given by:

$$A(n)=\exp FT_n,\quad T_n=t_{n+1}-t_n \qquad (1)$$

$$B(n)=\left(\int_0^{T_n}\exp F(T_n-s)ds\right)G_1 \qquad (2)$$

If the samples $\{u(t_n):n\in N_0\}$ are available, the sequence $\{u(h_n):n\in N_0\}$ cannot be found if no assumption on the class of admissible functions u(t) is imposed. In many practical situations, where u(t) is piecewise continuous, a set of k+1 previous samples $\{u(t_{n-k}),\ldots,u(t_{n-1}),u(t_n)\}$ can be used to evaluate $u(h_n)$ by extrapolation methods. In [18, pp. 39-52] a generalization of Newton-Gregory extrapolation [12] in the nonuniformly spaced data case is used for such evaluation. Also in [1, Ch. 2] and [30, Ch. 8] several interpolation methods by using spline functions are presented. State representation of interpolation algorithms [26] can be combined with the equation SD to include both the system representation and the extrapolation algorithm. For simplicity, the assumption that u(t) is piecewise constant ($u(t):t\in[t_n,t_{n+1}),\ n\in N_0$) is imposed. This assumption is not very crude; there are situations where this is true e.g. bang-bang control, pulse width modulation control, helsman's task [41] etc. The physical analog of such an assumption is to approximate a system S_1 by another one, preceeded by a sample and hold device. Then,

$$u(h_n)=u(t_n)\quad :n\in N_0 \qquad (3)$$

The sampled version of the stochastic model S_2 can be found by using an averaging sampler [39, pp. 174-176] as a sampling device. The result is a stochastic difference equation of the form:

$$x_2(t_{n+1}) = A(n)x_2(t_n) + w_2(n) \qquad (SS)$$

$$z(n) = C_2(n)x_2(t_n) + v_2(n)$$

where $z(n) = y_2(t_{r+1}) - y_2(t_n)$, $A(n)$ is given in (1), $C_2(n)$ is

$$C_2(n) = C_s \int_0^{T_n} \exp F(T_n - s)\,ds \qquad (4)$$

and the processes $\{w_2(n)\}$ and $\{v_2(n)\}$ are described statistically by:

$$E\{w_2(n)\} = 0; \quad \operatorname{cov}\{w_2(n), w_2(m)\} = Q_2(n)\delta_K(n-m) \qquad (5)$$

$$E\{v_2(n)\} = 0; \quad \operatorname{cov}\{v_2(n), v_2(m)\} = R_2(n)\delta_K(n-m) \qquad (6)$$

$$\operatorname{cov}\{w_2(n), v_2(m)\} = R_w(n)\delta_K(n-m) \qquad (7)$$

with,

$$Q_2(n) = \int_0^{T_n} A(T_n - s)G_2 Q G_2^T A^T(T_n - s)\,ds \qquad (8)$$

$$R_2(n) = C_s\left[\int_0^{T_n} B_2(T_n - s)QB_2^T(T_n - s)\,ds\right]C_s^T + T_n R \qquad (9)$$

$$R_w(n) = \left[\int_0^{T_n} A(T_n - s)G_2 Q B_2^T(T_n - s)\,ds\right]C_s^T \qquad (10)$$

$$A(T_n - s) = \exp F(T_n - s) \qquad (11)$$

$$B_2(T_n - s) = \left[\int_0^{T_n - s} A(T_n - r)\,dr\right]G_2 \qquad (12)$$

The proof of the above relationships can be found in [18, Ch. 2] and [3, pp. 82-86]. Eqs (5), (6) and (7) show that the processes $\{w_2(n)\}$ and $\{v_2(n)\}$ are nonstationary and correlated. The simulation of the system SS by a digital computer implies fast calculation of the integrals (8), (9), (10) and (12). This is discussed in Sec. III.

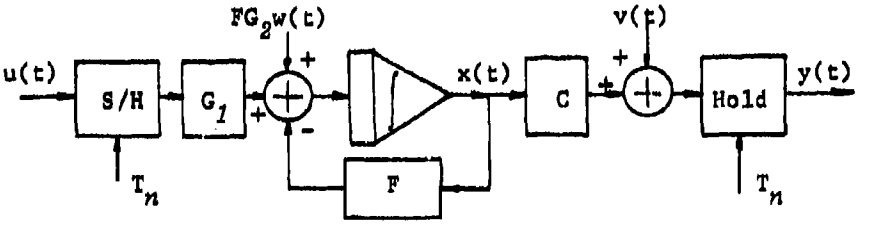


Fig. 4: The Model of the Plant

The conclusion of this section is that the plant S can be simulated by two systems SD and SS. The simulated states and outputs are:

$$x(t_n)=x_1(t_n)+x_2(t_n) \qquad (13)$$
$$y(t_n)=y_1(t_n)+y_2(t_n)$$

(The subscripts 1 and 2 denote the deterministic and the stochastic component of the variables respectively,) The overall model of the plant is illustrated in Fig. 4.

III. State Estimation of Nonuniformly Sampled Systems

Estimation theory is often employed to develop a memory representation of how the human gradually forms an internal model. In extensive studies on human decision making by Rouse [8, 33, 34], the subject is required to estimate the state of a dynamic process. For example, to successfully drive a car, purchase stocks and bonds, or plan a research project, the human must predict future states of various dynamic processes. The behavior of the human supervisor as a state estimator is widely used [16, 44].

Rouse's theory [34] on short-term and long-term adaptation is based on modeling the controlled plant as a time invariant discrete process and the human as a discrete estimator-identifier. This assumption can be crude in the following sense: visual scanning of the plant output display in general is not regular due to attention allocation caused by "control interrupts" and because of the nonconstant duration of the decision making process. Therefore the following idea is proposed: to use a discrete state estimator of the nonuniformly sampled version of the linear time invariant, dynamic plant. In the preceding Sec. III we saw that an appropriate model is not time invariant but time varying (Eqs. SD, SS, and (13)).

In Fig. 5 we relate the nonuniform estimator assertion to the model proposed by Kleinman et al. [16]. The vehicle dynamics, the motor noise and the observation noise are assumed to be part of the plant. For this section, we assume that the delays and the noise covariances $\{\tau_N, \tau, V_y, V_m\}$ are known. The objective is to design the equalization network or more specifically the estimator based on the sampling sequence $\{t_n : n \in N\}$. When V_y and V_m are given by Eqs. (21) and (22) of their paper [16], the adjustment of the model can be done by following their

procedure but with a different estimation algorithm.

In the following we state the estimation problem in the nonuniformly sampled data case, the difficulties involved and finally a possible solution.

We are given the Eq. S which is assumed to describe the plant dynamics, the neuromotor dynamics, the motor and the observation noise. Based on the sampling sequence $\{t_n : n \in N_0\}$(known) and the system S we can find the discrete equations SD, SS, and (13). The objective is to estimate the states $x(t_n)$ based on the observations $\{y(t_n) : n \in N\}$.

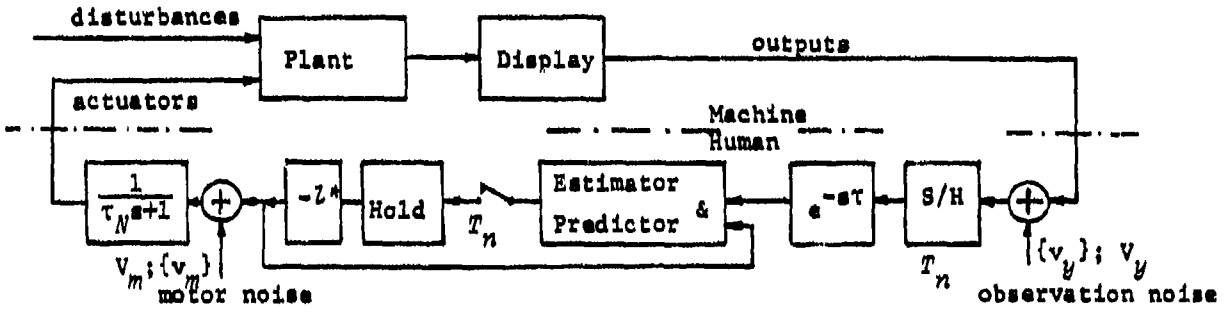


Fig. 5: Modifications of the Model of Human Operator

The extended Kalman filter [36, Ch. 8] can be applied for the above systems. The particular difficulties for such an application are:

- The driving noise $\{w_2(n)\}$ and the measurement noise $\{v_2(n)\}$ in SS are correlated according to (7) and (10).
- The overall estimator should contain the effect of the control input u(t).
- The estimator has to be updated from step to step because the system parameters A(n), B(n), $B_2(n)$, and $C_2(n)$ change from t_n to t_{n+1}; also the processes $\{w_2(n)\}$ and $\{v_2(n)\}$ are nonstationary ($R_2(n)$, $Q_2(n)$, and $R_w(n)$ are functions of T_n).

The first difficulty is overcome by transforming the system SS to another one with uncorrelated system and observation noises [36, Sec. 8.2]. Because of the linearity of both models SD and SS we can combine an open loop state observer for SD and a Kalman estimator for SS [39, Sec. 5-7]. In this manner the effect of the control input can be taken into account. The computational difficulties have been reduced to a minimum by two techniques: application of Sylvester's theorem and the use of Jordan canonical form. Both techniques are summarized in Appendix 1.

After the computational difficulties have been eliminated, a real time procedure can be applied for state estimation. The steps are presented below [18, pp. 184-200]:

Nonuniformly Sampled System State Estimation Algorithm

1. The algorithm starts with the following initial data

$$F,\ G_1,\ G_2,\ C,\ R,\ Q,\ \bar{x}_1(0),\ t_1-t_0,\ u(0),\ y(1),$$

and the following calculations

$$A(0),\ B(0),\ C_2(0),\ A_3(0),\ B_3(0),\ Q_2(0),\ R_2(0),\ R_w(0),\ Q_3(0),$$

(the formulas for the above evaluations are given in step 2),

$$\bar{x}_2(0)=0$$

$$\bar{x}_1(1)=A(0)\bar{x}_1(0)+B(0)u(0)$$

$$y_1(1)=C\bar{x}_1(1).$$

The next step is executed with n=0.

2. The algorithm continues using the $T_{n+1}=t_{n+2}-t_{n+1}$ interval.

$A(n+1)$ from Eq. (A2) or (A7),
$B(n+1)$ from Eq. (A3) or (A10),
$C_2(n+1)$ from Eq. (A3) or (A13),
$Q_2(n+1)$ from equation (A4),
$R_2(n+1)$ from Eq. (A5),
$R_w(n+1)$ from Eq. (A6),

$$A_2(n+1)=A(n+1)-R_w(n+1)R_2^{-1}(n+1)C_2(n+1) \tag{14}$$

$$B_3(n+1)=R_w(n+1)R_2^{-1}(n+1) \tag{15}$$

$$Q_3(n+1)=Q_2(n+1)-R_w(n+1)R_2^{-1}(n+1)R_w^T(n+1) \tag{16}$$

3. Computation of the prior estimates takes place. The input u(n+1) and output y(t+2) are sampled.

$$\bar{x}_1(n+2)=A(n+1)\bar{x}_1(n+1)+B(n+1)u(n+1)$$
$$y_1(n+2)=C\bar{x}_1(n+2)$$
$$z(n+1)=y(n+2)-\bar{y}_1(n+2)-[y(n+1)-\bar{y}_1(n+1)]$$
$$\bar{x}_2(n+1/n)=A_2(n)\bar{x}_2(n)+B_3(n)z(n)$$

$$P(n+1/n)=A_2(n)P(n/n)A_2^T(n) \tag{17}$$

4. The Kalman gain and estimate of x(n+1) are calculated now.

$$P(n+1/n+1)=[P^{-1}(n+1/n)+C_2^T(n+1)Q_3^{-1}(n+1)C_2(n+1)]^{-1}P(n+1/n) \tag{18}$$

$$K(n)=P(n+1/n+1)C_2^T(n+1)Q_3^{-1}(n+1)$$

$$\bar{x}_2(n+1)=\bar{x}_2(n+1/n)+K(n)[z(n)-C_2(n+1)\bar{x}_2(n+1/n)] \tag{19}$$

$$\bar{x}(n+1)=\bar{x}_1(n+1)+\bar{x}_2(n+1)$$

($\bar{x}(n+1)$ is the estimate of $x(t_{n+1})$). The algorithm continues at step 2.

The major result of this section is the previous algorithm with "input" the continuous time parameters (F, G_1, G_2, C, R and Q) and "output" the estimate $\bar{x}(n+1)$. The intervals T_n have no restrictions (they can be arbitrarily large or small) but they must known explicitly. For theoretical studies a Prediction Error Model [11] or an Associated Model [44] of the above estimation algorithm is available. The form is:

$$x(n+1)=f_1[\,x(n),u(n),n\,]+f_2(n)v(n) \tag{20}$$

$$y(n)\;\;=g_1[\,x(n),u(n),n\,]+v(n)$$

where $\{v(n)\}$ is a white Gaussian, zero mean process (innovations). The functions $f_1(.)$, $f_2(.)$, $g_1(.)$ and the covariance of $\{v(n)\}$ are calculated recursively [18, Sec. 4-2]. The delay τ (Fig. 5) can be taken into account at the forth step of the algorithm by adding the equation

$$x(t_n+\tau)=A(\tau)x(n).$$

Obviously, the structure of the optimal controller (the gain l^*) is not affected by using the previous algorithm.

IV. Parameter Identification Using Irregularly Spaced Samples

Having fixed the overall description of the controlled plant, the parameters of the human model can be found by an identification scheme. In many other cases the model of the plant is not exactly known and a composite man-machine model is desired. Generalizing the identification procedures in the nonuniformly sampled data case we look for a model of the man-machine task. The statement of the problem follows.

We are given the structure of a system S_p (which represents the controlled

plant and the human operator) in terms of a vector of unknown parameters p. Mathematically the system is assumed to be described by:

$$dx=F_s(p)xdt+G_{1s}(p)udt+G_{2s}dw \qquad (S_p)$$
$$dy=C_s(p)xdt+D_s(p)udt+dv$$

where p might contain unknown elements of F_s, G_{1s}, C_s. Because of the nature of the system or because of our inability to observe the system regularly, we assume that the irregularly spaced samples of u(t) and y(t) at $t=t_n; n\epsilon N$ are given.
The objective is to find an estimate of the vector p.

Additional assumptions are imposed on the above system.

1. The noises {v(t)} and {w(t)} are described statistically as Wiener-Levy processes with known covariances Q and R.
2. The initial state x(0) of the system S_p and the vector p are assumed to be Gaussian random vectors with means x_o and p_o and variances S_x and S_p respectively. This restriction is not severe; for example, if we let $S_p=\sigma I$ with σ a large number we do include a large class of practical problems with "completely unknown" p.
3. Identifiability [4] and resolvability [17] criteria should be satisfied. Improper initial states and/or system representation might result in an unidentifiable system. For the following development we assume that any system $\{F_s', G_{1s}', C_s'\}$ algebraically equivalent to $\{F_s, G_{1s}, C_s\}$ is a solution of the problem (actually in [17] it is stated that identification of the actual system S_p cannot be done but only identification of the associated system (20)).
4. The sampling sequence is not unrestricted. In [18, pp. 133-149] it is shown that there exist sequences $\{t_n\}$ under which the sampled version of S_p is unobservable even if S_p is observable; thus, ill conditioned sequences must be excepted from the identification procedure.

The parameter identification problem is solved by the following two methods.

First Approach: Linear Estimation Formulation

In this subsection, we deal with a particular class of nonuniformly sampled systems. The state representation of the linear system before sampling is assumed to be (Pearson [28]):

$$\dot{x}(t)=F(p)x(t)+G(p)u(t) \qquad (21)$$
$$y(t)=Cx(t)$$
$$x(0)=x_o; \quad u\epsilon R^{\ell}; \; x\epsilon R^{mq}; \quad y\epsilon R^{m}$$

$$F(p)=\begin{bmatrix} -F_1 & I & 0 & \cdots & 0 \\ -F_2 & 0 & I & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ -F_q & 0 & 0 & \cdots & I \end{bmatrix} \qquad G(p)=\begin{bmatrix} G_1 \\ G_2 \\ \cdots \\ G_q \end{bmatrix} \qquad C=[\, I_m \; 0 \; 0 \; \cdots \; 0 \,] \tag{22}$$

where F is qmxqm, F_i is mxm, G is qmxℓ, G_i is mxℓ; C is mxqm and p is a vector containing all the unknown elements of F and G. The identification problem is to find the vector p when the measurements $y(t_n)$ and the controls $u(t_n)$ are available.

Let us use the notation:

- $\bar{F}_i$, $\bar{G}_i$, $\bar{x}(t)$, $\bar{y}(t)$ to stand for the estimates of F_i, G_i, x, and y respectively.
- DF_i, DG_i to mean the errors $\bar{F}_i-F_i$ and $\bar{G}_i-G_i$ respectively.
- e to be the error of the estimates of p, defined by:

$$e=\begin{bmatrix} Df \\ Dg \end{bmatrix}; \qquad Df=[\, DF_{11} \; DF_{12} \cdots DF_{qm} \,]^T; \quad Dg=[\, DG_{11} \; DG_{12} \cdots DG_{q\ell} \,]^T \tag{23}$$

where DF_{ij} and DG_{ij} are the j-th rows of DF_i and DG_i respectively (Df is m^2q-vector, and Dg is $mq\ell$-vector).

- Y(t) and U(t) to mean the following matrices:

$$Y(t)=\begin{bmatrix} y^T(t) & 0 & \cdots & 0 \\ 0 & y^T(t) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & y^T(t) \end{bmatrix} \qquad U(t)=\begin{bmatrix} u^T(t) & 0 & \cdots & 0 \\ 0 & u^T(t) & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & u^T(t) \end{bmatrix} \tag{24}$$

It can be proved [18, pp.277-288] that the above quantities satisfy the following (linear) equation:

$$y(t_n)-\bar{y}(t_n)=H(t_n)e+\omega(t_n) \tag{25}$$

$$H(t_n)=C\int_0^{t_n} \exp F(t_n-s)\;[Y(s)\mid U(s)]ds \tag{26}$$

$$\omega(t_n)=C\exp F(t_n-t_o)\;[x(t_o)-\bar{x}(t_o)]$$

At this point the parameter estimation problem is considered to be solved. Eq. (25) is a linear set of equations of the unknown errors e; the function $H(t_n)$ can be calculated from input-output signals. Notice also that the term $\omega(t)$ vanishes quickly if the eigenvalues of the selected matrix $\bar{F}$ are located in the left half plane. Evaluation of H(.) can be done by the following methods:

- Using measurement devices.
- Using a computer extrapolation algorithm.

The first method becomes clearer by viewing H(.) as an output (matrix) of a set of linear systems starting from zero initial state (see also Fig. 6).

$$\dot{w}_1(t)=\bar{F}w_1(t)+Y(t) \qquad \text{output measurement system} \qquad (27)$$
$$w_1(t_0)=0;$$

$$\dot{w}_2(t)=\bar{F}w_2(t)+U(t) \qquad \text{input measurement system} \qquad (28)$$
$$w_2(t_0)=0,$$

and

$$H(t_n)=C[w_1(t_n) \mid w_2(t_n)] \qquad (29)$$

The second method is very similar to the previous one; digital filters can be implemented to interpolate the points $y(t_n)$ and $u(t_n)$ and evaluate the integral in (26) by using a computer algorithm. After $H(t_n)$ has been found, estimation of the error e of the parameter p can be done by:

- the least squares method (Sandoz [35], Hsia [14]),
- an adjustable observer (Luenberger [21, 22], Narendra [23]),
- or a recursive linear, time varying Kalman filter.

If the measurements are noisy it is known that the least squares method give a biased estimator. Several algorithms exist to remove the bias error term. (See Landau [19], and Astrom [4].) We intend to consider this identification scheme as a generalization of the instrumental variables method (Young [45, 46]) in the non-uniformly sampled data case. Gabav and Merhav [10, 9] have successfully applied Young's method to identify human operator models; using the modifications we suggest the above scheme can be generalized in the nonuniformly sampled data case.

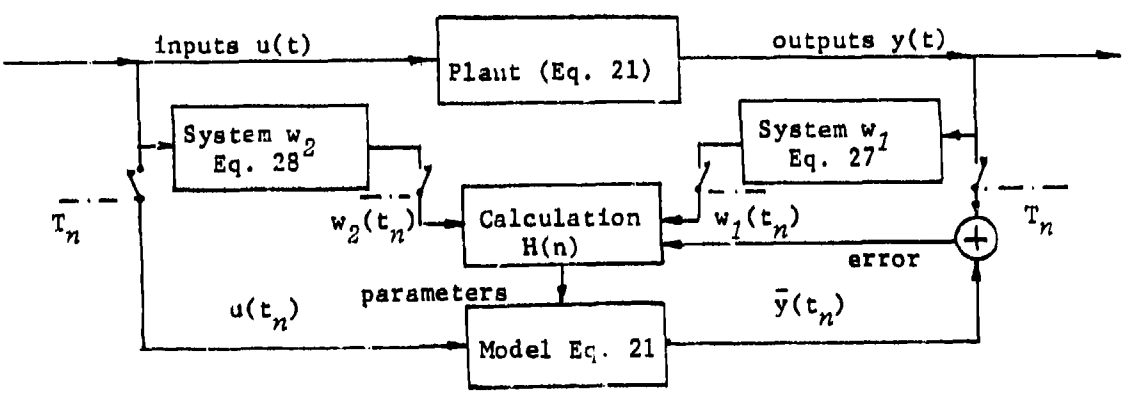


Fig. 6: Parameter Identification Using Measurement Devices

Second Approach: Nonlinear Estimation Formulation

The previous approach has a disadvantage: it does not take into account the statistics of the system and measurement noise; the method is primarily useful for systems with low level noise. In this subsection we outline a more general approach and then we present it.

Step 1: Formulation of the parameter identification problem as a nonlinear, time varying, state estimation problem of an augmented system is accomplished.

Step 2: Application of the maximum likelihood, the maximum aposteriori and the weighted least squares method results in a deterministic (cost) function; minimization of such a function subject to system constraints assure us that the model matches to the input and output samples.

Step 3: Minimization of the cost function is accomplished by using variational calculus. The result is a two-point boundary-value problem (TPBVP).

Step 4: We present an off-line algorithm that can be used to solve the TPBVP. Several other algorithms can be used for the same purpose.

Step 5: We present an on-line algorithm that can be used to solve the TPBVP real time. The algorithm is based on the invariant imbedding equations.

Because of the length of the proofs most of them have been omitted; appropriate references are given to direct the reader.

Step 1: The Augmented System

The sampled version of equations S_p can be found by using the results of Sec. II:

$$x_1(t_{n+1}) = A(p;T_n)x_1(t_n) + B(p;T_n)u(t_n) \qquad (SD_p)$$

$$y_1(t_n) = C(p)x_1(t_n)$$

$$x_2(t_{n+1}) = A(p;T_n)x_2(t_n) + w_2(t_n) \qquad (SS_p)$$

$$y_2(t_{n+1}) - y_2(t_n) = C_2(p;T_n)x_2(t_n) + v_2(t_n)$$

$$p(n+1) = p(n)$$

The last equation means that the unknown parameters of F, G_1, and C should remain invariant with respect to time. Parameter identification of systems SD_p and SS_p can be seen as state estimation of an augmented system of the form:

$$x_a(n+1) = A_a[x_a(n),n] + B_a(n)w_a(n) \qquad (S_a)$$

$$y_a(n) = C_a[x_a(n),n] + v_a(n),$$

where

$$x_a(n) = \mathrm{col}[x_1(t_n),\ x_2(t_n),\ p(n)]; \quad y_a(n) = y(t_{n+1}) - y(t_n)$$

and the functions A [.], B [.], and C [.] are defined properly in terms of $A(p;T_n)$, $B(p;T_n)$, and $C_2(p;T_n)$ (Schweppe [39], Sage and Melsa [38], [18, pp. 211-215]).

Step 2: The Cost Function

Application of the maximum likelihood method to determine an estimate $\bar{x}_a(n)$ of the augmented state $x_a(n)$ of S_a requires the following assumptions:

1. $\{v_a(n)\}$ to be white Gaussian noise with zero mean and variance $R_a[x_a(n),n]$.
2. $\{w_a(n)\}$ to be a stochastic process with zero mean and uncorrelated to $\{v_a(n)\}$.
3. The observations $\{y(t_1), y(t_2), \ldots y(t_N)\}$ to be available.

Then, maximization of the conditional probability density:

$$p[y_a(t_1),\ldots,y_a(t_N)/x_a(t_1),\ldots,x_a(t_N)]$$

is equivalent to minimization of the cost function

$$JML=1/2 \sum_{j=0}^{N-1} \| y_a(t_{j+1})-C_a[x_a(j+1),j+1]; R_a^{-1}[x_a(j+1),j+1] \|^2+\ln\det R_a[x_a(j+1),j+1] \quad (31)$$

(the notation $\| z; Q\|^2$ means $z^T Qz$)

with respect to $\{x_a(1),\ldots,x_a(N)\}$ and subject to the equality constraints

$$x_a(j+1)=A_a[x_a(j),j] \qquad j=1,2,\ldots,N-1 \quad (32)$$

In the same manner, application of the maximum aposteriori method requires the following additinal assumptions:

4. $\{w_a(n)\}$ to be white Gaussian process with zero mean and variance $Q_a[x_a(n),n]$.
5. $x_a(t_o)$ to be a Gaussian random vector with mean x_o and variance S_a.

Then, maximization of the aposteriori probability density:

$$p[x_a(t_o),\ldots,x_a(t_N)/y_a(t_1),\ldots,y_a(t_N)]$$

is equivalent to minimization of the cost function JMA:

$$JMA=q_0+q_1+q_2+q_3 \quad (33)$$

$$q_0=1/2\|x_a(t_o)-x_o; S_a^{-1}\|^2 \quad (34)$$

$$q_1=1/2 \sum_{j=0}^{N-1} \| y_a(j+1)-C_a[x_a(j+1),j+1]; R_a^{-1}[x_a(j+1),j+1] \|^2 \quad (35)$$

$$q_2=1/2 \sum_{j=0}^{N-1} \| \bar{w}_a(j); Q_a^{-1}[x_a(j),j] \|^2 \quad (36)$$

$$q_3=1/2 \sum_{j=1}^{N} \ln\det R_a[x_a(j),j]\det Q_a[x_a(j-1),j-1] \quad (37)$$

subject to the system equalities (32) and using

$$\bar{w}_a(j)=x_a(j+1)-A_a[x_a(j),j] \tag{38}$$

In order to apply the least squares method to minimize the initial state error, the output error and the trajectory error between the real world samples and the output of the model S_a, we need the following cost function:

(39)

$$JLS=1/2\| x_a(0)-x_o;S_a^{-1}\|^2+1/2\sum_{j=0}^{N-1}\| y_a(j+1)-C_a[x_a(j+1),j+1];R_a^{-1}(j+1)\|^2+\| w_a(j);Q_a^{-1}(j)\|^2$$

where S_a, R_a and Q_a are given (apriori) positive semidefinite matrices which do not depend on x_a.

The proof of the above statements ((31), (33) and (39)) can be found in [18, pp. 215-237, also in 39 and 33]. For the following development we use the function JLS to derive the TPBVP; if the assumptions under which minimization of (39) does not correspond to exact model matching are not met, JML or JMA can be used. The following steps do apply to more general cost functions.

Step 3: The TPBVP

The trajectory of the state x_a that minimizes the cost function JLS can be found by using the maximum principle [5, 37]. Let us simplify the notation (and make our procedure more general) by using:

$$\phi[x_a(i),\bar{w}_a(i)]=A_a[x_a(i),i]+B_a(i)\bar{w}_a(i) \tag{40}$$

$$\theta_o[x_a(0)]=1/2\,\|x_a(0)-x_o;\ S_a^{-1}\|^2 \tag{41}$$

$$\theta[x_a(i),\bar{w}_a(i)]=1/2\,\|y_a(i+1)-C_a[\phi[x_a(i),\bar{w}_a(i)],i+1];\ R_a^{-1}(i+1)\|^2+1/2\,\|w_a(i);\ Q_a^{-1}\|^2 \tag{42}$$

$$X_N=\begin{bmatrix}x_a(0)\\ \vdots\\ x_a(N)\end{bmatrix}\qquad W_N=\begin{bmatrix}w_a(0)\\ \vdots\\ w_a(N)\end{bmatrix} \tag{43}$$

The necessary conditions for minimization of

$$JLS(X_N,W_N)=\theta_o[x_a(0)]+\sum_{k=0}^{N-1}\theta[x_a(k),\bar{w}_a(k)] \tag{44}$$

subject to the system equalities are:

$$\partial H/\partial x_a(n)=\lambda(n);\quad \partial H/\partial\lambda(n+1)=x_a(n+1);\quad \partial H/\partial\bar{w}_a(n)=0 \tag{45}$$

where H(.) is the Hamiltonian

$$H[x_a(i),\lambda(i),\bar{w}_a(i)]=\theta[x_a(i),\bar{w}_a(i)]+\lambda^T(i+1)\,\phi[x_a(i),\bar{w}_i(i)] \tag{46}$$

and $\lambda(i+1)$ is a vector (costate) of Lagrange multipliers associated with $x_a(i)$. The result of the minimization is the following TPBVP:

$$x_a(n+1)=A_a[x_a(n),n]-B_a(n)Q_a(n)B_a^T(n)A_a^{*-1}\lambda(n) \tag{47}$$

$$\lambda(n+1) =A_a^{*-1}\lambda(n)+C_a^*R_a^{-1}(n+1)\ [y_a(t_{n+1})-C_a[x_a(n+1),n+1]] \tag{48}$$

$$\lambda(0)=S_a^{-1}[x_o-x_a(0)] \tag{49}$$

$$\lambda(N)=0$$

where,

$$A_a^* =\partial A_a^T[x_a(n),n]/\partial x_a(n) \tag{50}$$

$$C_a^* =\partial C_a^T[x_a(n+1),n+1]/\partial x_a(n+1).$$

Observe that if the dimension of the vector x_a is n_a, Eqs (47), (48) and (49) consist a system of $(2N+1)n_a$ nonlinear equations with $2N+1$ times n_a unknowns. Also, the above result is valid under the assumptions of JLS. If JML or JMA are desired to be used, equations similar to (47) and (49) can be derived.

Step 4: Off-line Solution of the TPBVP

In the literature [29] there exist several numerical methods to solve the nonlinear TPBVP. For example, first order gradient (steepest descent), second order gradient, conjugate gradient and quasilinearization methods can be used. In the following we present a modification of the first order gradient method presented in [38]. The basic idea behind the method is to satisfy the set of equations:

$$\partial H/\partial\lambda(k+1)=x_a(k+1);\quad \partial H/\partial x_a(k)=\lambda(k);\quad \partial H/\partial\bar{w}_a(k)=0 \tag{51}$$

and to try iteratively to satisfy the boundary conditions (49). In [18] it is proved that if (51) are satisfied the first variation of JLS (δJLS) is given by:

$$\delta JLS=\left[\frac{d\theta_o^T[x_a(0)]}{dx_a(0)}\ \vdots\ -\lambda(N)\right]\begin{bmatrix}x_a(0)\\ x_a(N)\end{bmatrix} \tag{52}$$

Application of the steepest descent method gives that the proper steps $\Delta x_a(0)$ and $\Delta x_a(N)$ should be

$$\Delta x_a(0)= -k_x\frac{d\theta_o[x_a(0)]}{dx_a(0)}\ ;\qquad x_a(N)= k_\lambda\lambda(N) \tag{53}$$

where k_x and k_λ are properly selected positive numbers. The algorithm proceeds as follows.

Estimation and Identification Algorithm

1. Pick a vector $x_a(N;i)$. (The argument i means iteration step.)
2. Find the vectors $x_a(0;i)$ and $\lambda(0;i)$ by solving the equations (17), and (18) backwards and by using initial states $\lambda(N;i)=0$ and $x_a(N;i)$ from step 1.
3. Update $x_a(0;i)$ and $\lambda(0;i)$ by using

$$x_a(0;i+1)=x_a(0;i)-k_x\frac{d\theta_o[x_a(0;i)]}{dx_a(0;i)}$$

$$\lambda(0;i+1)=\lambda(0;i)$$

4. Find the vectors $x_a(N;i+1)$ and $\lambda(N;i+1)$ by solving the equations (17) and (18) forwards and by using initial states $\lambda(0;i+1)$ and $x_a(0;i+1)$.
5. Update $x_a(N;i+1)$ and $\lambda(N;i+1)$ by:

$$x_a(N;i+2)=x_a(N;i+1)-k_\lambda\lambda(N;i+1)$$

$$\lambda(N;i+2)=0$$

6. Continue to step 2 until there is no significant change of $x_a(0)$ and $\lambda(N)$.

The major advantage of this algorithm versus the existing ones is the low memory requirements. Array processors promise a reasonable computation time for solving problems of the above form.

Step 5: Sequential solution of the TPBVP.

The sequential solution of the nonlinear, time varying, state estimation (and parameter identification problem) is a current area of research. A brief review of the available techniques follows.

- Linearization of the Kalman filter equations about a nominal point. This method is based on the Kalman filter as an estimator of states and parameters after linearization of the functions $A_a[x_a(n),n]$ and $C_a[x_a(n),n]$ about a nominal point $x_m(t_n)$. The major disadvantage of the method is the selection of the nominal point [39].
- Linearization of the Kalman filter about the old estimate. The method uses as nominal point the previous estimate of $x_a(n)$. Serious stability problems might arise [39, 36].
- Selection of a linear model out of a set of precalculated models (partitioned algorithms). The basic idea of the method is to linearize the model about several

predetermined operating points and then use a Kalman filter for each operating point as a state estimator. A decision rule can be used afterwards to select the model that gives the best results. Extensive literature on this approach can be found in [20, 24].

- Linearization of the invariant imbedding equations. This technique is widely used to solve the TPBVP in a sequential manner [26, 5, 37, 38]. The algorithm is based on the following equations [18, pp. 291-301];

One step predictor,

$$x_\alpha(n+1/n)=A_\alpha[x_\alpha(n),n]$$

Filter algorithm,

$$x_\alpha(n+1)=x_\alpha(n+1/n)+P(n+1)C_\alpha^{*}R_\alpha^{-1}(n+1)\ [y_\alpha(n+1)-C_\alpha[x_\alpha(n+1/n),n+1]]$$

$$C_\alpha^{*}=\partial C_\alpha[w,n+1]/\partial w; \qquad w=x_\alpha(n+1/n).$$

Prior variance.

$$P(n+1/n)=A_\alpha^{*T}P(n)A_\alpha^{*}+B_\alpha(n)Q_\alpha(n)B_\alpha^{T}(n)$$

$$A_\alpha^{*}=\partial A_\alpha[w,n]/\partial w; \qquad w=x_\alpha(n).$$

Error variance,

$$P(n+1)=[I+M_\alpha^{*T}P(n+1/n)]^{-1}P(n+1/n)$$

$$M_\alpha^{*}=\frac{\partial}{\partial w}C_\alpha^{T}(w,n+1)R_\alpha^{-1}(n+1)C_\alpha^{*T}(w,n+1); \qquad w=x_\alpha(n+1/n)$$

Initial conditions:

$$x_\alpha(0)=x_o; \qquad P(0)=S_\alpha$$

Because the functions $A_\alpha(.)$, and $C_\alpha(.)$ with respect to $x_\alpha^{T}=(x_1^{T}\ x_2^{T}\ p^{T})$ are known, the gradiant functions can be calculated analytically. In [18] the above calculations are presented when the Jordan diagonal form is assumed and single input systems are considered.

V. Conclusion

Several resons for using nonuniformly sampled systems as models for human motivated this study. A comprehensive theory of the irregular sampling process is the starting point of the whole development; then, the nonuniformly sampled models are used for solving two fundamental problems: the state estimation and the parameter identification. Applicability of the results relies on the computational studies of the presented algorithms.

There are several trends and extentions of this research. Practical manual

control problems are only a subclass of the applications of the theory. Each individual case requires special treatment and futher study. One direction of our work is towards solution of the nonlinear estimation problem real-time by using partitioned algorithms; another, is towards solution of the estimation problem in the case of statistically known intervals. Sampling sequences with given stochastic characteristics need futher study.

Appendix 1: Computational Aspects
(from Sec. III)

The first computational technique is based on the following procedure: At the beginning, the component matrices [48] $F_i: i=1,2,\ldots,n$ of F are calculated and stored. Depending on the multiplicity of the eigenvalues of F, Sylvester's theorem or the extension of Sylvester's theorem [18] is applied to find that

$$g(F)=F_1 g(s_1)+\ldots+F_n g(s_n) \tag{A1}$$

where g(.) is an analytic function, and s_i are the eigenvalues of F. The transition matrix A(n), and the integrals involved in the calculation of B(n), $C_2(n)$. $R_2(n)$, $Q_2(n)$ and $R_w(n)$ can be evaluated by properly changing the definition of the g(.) function. The results are:

$$A(n) = \sum_i F_i g_1(s_i) \tag{A2}$$

$$B(n) = [\sum_i F_i g_2(s_i)]G_1; \qquad C_2(n)=C_s[\sum_i F_i g_2(s_i)] \tag{A3}$$

$$Q_2(n)=\sum_i\sum_j F_i G_2 Q\, {}^T_2 F_j g_3(s_i,s_j) \tag{A4}$$

$$R_2(n)=C[\sum_i\sum_j F_i G_2 Q G_2^T\, F_j g_4(s_i,s_j)]C^T+T_n R \tag{A5}$$

$$R_w(n)=[\sum_i\sum_j F_i G_2 Q G_2^T\, F_j g_5(s_i,s_j)]C^T \tag{A6}$$

$$(i,j)\in\{1,2,\ldots,n\}^2$$

where $g_j(.)$ $j=1,2,\ldots,5$ are appropriate functions with parameter T_n ([18] pp. 102-120). Observe that calculation of the above parameters from t_n to t_{n+1} involves a summation of n or n^2 terms; numerical integration is completely avoided.

The second method is based on the Jordan form representation of the system S [6 Sec. 2-6]. The calculation of the system parameters A(n), B(n) and C (n) is based on the following equations:

$$A(n)=\mathrm{diag}[A_1(n)\ldots A_r(n)] \qquad n\times n \tag{A7}$$

$$A_i(n)=\mathrm{diag}[A_{i1}(n)\ldots A_{ir_i}(n)] \qquad n_i\times n_i \tag{A8}$$

$$A_{ij}(n)=M_{ij}(T_n)\exp s_i T_n \qquad n_{ij}\times n_{ij} \tag{A9}$$

$$B(n)=\mathrm{col}[B_1(n)\ldots B_r(n)] \qquad l\times n \tag{A10}$$

$$B_i(n)=\mathrm{col}[B_{i1}(n)\ldots B_{ir_i}(n)] \qquad l\times n_i \tag{A11}$$

$$B_{ij}(n)=\begin{cases}[Z_i(T_n)\exp s_i T_n - f_i I]G_{ij}/s_i & \text{if } s_i\neq 0\\ M_{ij}^{(-1)}(T_n)G_{ij} & \text{if } s_i=0\end{cases} \tag{A12}$$

$$C_2(n)=\text{row}[c_1(n)\ldots c_r(n)] \qquad m\times n \tag{A13}$$

$$c_i(n)=\text{row}[C_{i1}(n)\ldots C_{ir_i}(n)] \qquad m\times n_i \tag{A14}$$

$$C_{ij}(n)=\begin{cases}C_s[Z_i(T_n)\exp s_i T_n - f_i I]/s_i & \text{if } s_i\neq 0\\ C_s M_{ij}^{(-1)}(T) & \text{if } s_j=0\end{cases} \tag{A15}$$

where $M_{ij}(T_n)$, $M_{ij}^{(-1)}(T_n)$ and $Z_i(T_n)$ are matrices with terms of the form $T_n^k/k!$ and,

$$f_i=1+1/s_i+\ldots+1/s_i^{n_{ij}} \qquad [18]$$

(r is the number of distinct eigenvalues of F, r_i is the number of Jordan blocks associated with the eigenvalue s_i, n_i is the dimension of the joint Jordan blocks associated with the eigenvalue s_i and n_{ij} is the dimension of the j-th Jordan block).

The computation time of the above parameters is very low and the problem of evaluating A(n), B(n), $C_2(n)$ for various n has no difficulty. Evaluation of $Q_2(n)$, $R_2(n)$ and $R_w(n)$ by following the same technique is also possible [18, pp. 130-134].

References

[1] Ahlberg J.H., E.N. Nilson and J.L. Walsh, The Theory of Splines and Their Applications. Academic Press, Inc., 1967.

[2] Allen R.W. and D.T. McRuer, "Driver Steering Dynamic Measured in a Car Simulator Under a Range of Visibility and Roadmaking Conditions," Proc, 13th Annual Conf. on Manual Control, pp. 180-199, June 1977.

[3] Astrom Karl J., Introduction to Stochastic Control Theory. New York: Academic Press, Inc., 1970.

[4] Astrom K.J., and P. Evkhoff, "System Identification. A Survey," Automatica, vol. 7, pp. 123-162, 1971.

[5] Bellman Richard, Introduction to the Mathematical Theory of Control Processes. New York: Academic Press, Inc., vol. I & II, 1967.

[6] Chen C.T., Introduction to Linear System Theory. Holt, Rinehard and Winston, Inc., 1970.

[7] Conant R.C. and W.R. Ashby, "Every good Regulator of a System must be a Model of that System," International Journal of Systems Science, vol. 1, no. 2, pp. 89-97, 1970.

[8] Enstrom Kenneth D. and William B. Rouse, "Real Time Determination of How a Human Has Allocated His Attention Between Control and Monitoring Tasks," IEEE Trans. on Systems, Man and Cybernetics, vol. SMC-7, no. 3, pp. 153-161, March 1977.

[9] Gabay E. and S.J. Merhav, "Identification of Linear Systems with Time Delay Operating in a Closed Loop in the Presence of Noise," IEEE Trans. on Autom. Control, Oct. 1976.

[10] _____, "Identification of a Parametric Model of the Human Operator in Closed-Loop Control Tasks," IEEE Trans. on Systems, Man and Cybernetics, vol. SMC-7, no. 4, pp. 284-292, April 1977.

[11] Goodwin Graham C. and Robert L. Payne, Dynamic System Identification: Experiment Design and Data Analysis. New York: Academic Press, Inc., 1977.

[12] Hamza M.H. and M.A. Sheirah, "Polynomial Extrapolation in Sampled Data Control Systems," IEEE Trans. Autom. Control, vol. AC-16, pp. 102-104, Feb. 1971.

[13] Harrington Walter W., "An Approach to Multi-Axis Proolem in Manual Control," Proc. 13th Annual Conf. on Manual Control, pp. 58-71, June 1977.

[14] Hsia T.C., System Identification, Massachusetts: Lexington Books, 1977.

[15] Kvalseth Tarald O., "A Decision-Theoretic Model of the Sampling Behaviour of the Human Process Monitor," IEEE Trans. of Systems, Man and Cybernectics, vol. SMC-7, no. 11, pp. 810-812, Nov. 1977.

[16] Kleinman D.L., S. Baron and W.H. Levison, "A Control Theoretic Approach to Manned-Vehicle Systems Analysis," IEEE Trans. on Autom. Control, vol. AC-16, Dec. 1971.

[17] Kok J.J. and R.A. Wijk, "A Model of the Human Supervisor," Proc. 13th Annual Conf. on Manual Control, pp. 210-217, June 1977.

[18] Kontopidis G.D., Nonuniformly Sampled Systems. NSEE Thesis, University of New Hampshire, Dec. 1978.

[19] Landau I.D., "Unbiased Recursive Identification Using Model Reference Adaptive Techniques," IEEE Trans. on Autom. Control, vol. AC-21, pp. 194-200, April 1976.

[20] Lainiotis Demetrios G., "Optimal Adaptive Estimation: Structure and Parameter Adaptation," IEEE Trans. on Autom. Control, vol. AC-16 no. 2, pp. 160-170, April 1972.

[21] Luenberger D.G., "Observing the State of Linear System," IEEE Trans. Mil. Electron., vol. MIL-8, pp. 74-80, April 1968.

[22] _____, Optimization by Vector Space Methods. New York: John Wiley & Sons, Inc., 1969.

[23] Luders Gerd, Kumpati S. Narendra, "An Adaptive Observer and Identifier for a Linear System," IEEE Trans. Autom. Control, vol. AC-17, pp. 496-499, Oct. 1973.

[24] Magill D.T., "Optimal Adaptive Estimation of Sampled Stochastic Processes," IEEE Trans. on Autom. Control, vol. AC-10, no. 4, pp. 434-439, Oct. 1965.

[25] McRuer D. and D.H. Weir, "Theory of Manual Vehicular Controls," IEEE Trans. Man-Machine Systems, vol. MMS-10, no. 4, pp. 257-291, 1969.

[26] Morrison N., Introduction to Sequential Smoothing and Prediction. New York: McGraw-Hill, 1969.

[27] Palmer E., "Interrupted Monitoring of a Stochastic Process," Proc. 13th Annual Conf. on Manual Control, pp. 237-245, June 1977.

[28] Pearson A.E., "Finite Time Interval Linear System Identification Without Initial State Estimation," Automatica, vol. 12, pp. 577-587, 1976.

[29] Polak E., Computational Methods in Optimization. New York: Academic Press, Inc., 1976.

[30] Pressman R.S. and J.E. Williams, Numerical Control and Computer Aided Manufacturing. New Jersey: John Wiley and Son, 1977.

[31] Repperger D.W. and E.J. Hartzell, "A Relationship Between Eye Movement Patterns and Performance in a Precognitive Tracking Task," Proc. 13th Annual Conf. on Manual Control, pp. 152-165, June 1977.

[32] Repa Brian S. and Robert S. Zucker, "The Application of Integral Performance Criteria to the Analysis of Discrete Maneuvers in a Driving Simulator," Proc. 13th Annual Conf. on Manual Control, pp. 81-100, June 1977.

[33] Rouse W.B., "Cognitive Sources of Suboptimal Human Prediction," Ph.D. Dissertation, M.I.T., 1972.

[34] _____, "A Theory of Human Decision Making in Stochastic Estimation Tasks," IEEE Trans. on Systems, Man and Cybernetics, vol. SMC-7, no. 4, pp. 274-283, April 1977.

[35] Sandoz D.J. and B.H. Swanick, "A Recursive Least Squares Approach to the Adaptive Control Problem," Int. J. Control, vol. 16, pp. 243-258, 1972.

[36] Sage Andrew P. and James Melsa, Estimation Theory with Applications to Communications and Control. New York: McGraw-Hill, 1971.

[37] _____ and Chelsea C. White, Optimum Systems Control, second edition, Englewood Cliffs, N.J.: Prentice Hall, 1977.

[38] _____ and James L. Melsa, System Identification. New York: Academic Press, Inc., 1971.

[39] Schweppe Fred C., Uncertain Dynamic Systems. Englewood Cliffs, N.J.: Prentice Hall, 1973.

[40] Tulga M.K. and T.B. Sheridan, "Supervisory Dynamic Decision Making in Multi-Task Monitoring and Control," Proc. 13th Annual Conf. on Manual Control, pp. 199-209, June 1977.

[41] Veldhuyzen W., "Modeling the Human Operator of Slowly Responding Systems Using Linear Models," Proc. 13th Annual Conf. on Manual Control, pp. 172-178, June 1977.

[42] Verplank W.L., "The Facilitating Effects of Uncertainty in Long Term Manual Control," Proc. 13th Annual Conf. on Manual Control, pp. 101-106, June 1977.

[43] Wewerinke P.H., "Performance and Workload Analysis of an In-Flight Helicopter Tasks," Proc. 13th Annual Conf. on Manual Control, pp. 106-118, June 1977.

[44] Wijk Ron A. and Jan J. Kok, "Theoretic Aspects of the Identification of Parameters in the Optimal Control Model," Proc. 13th Annual Conf. on Manual Control, pp. 27-35, June 1977.

[45] Young P.C., "An Instrumental Variable Method for Real Time Identification of a Noisy Process," Automatica, vol. 6, pp. 271-287, 1970.

[46] _____, "Comments on Line Identification of Linear Dynamic Systems with Applications to Kalman Filtering," IEEE Trans. Automa. Control, vol. AC-17, pp. 269-270, April 1972.

[47] Zacharias G.L. and L.R. Young, "Manual Control of Yaw Motion with Combined Visual and Vestibular Cues," Proc. 13th Annual Conf. on Manual Control, pp. 389-403, June 1977.

[48] Zadeh Lotfi A. and Charles A. Desoer, Linear System Theory: The State Space Approach. New York: McGraw-Hill, 1963.

DIVISION OF ATTENTION AS A FUNCTION OF THE NUMBER OF STEPS, VISUAL SHIFTS, AND MEMORY LOAD*

Richard A. Chechile

Keith Butler

William Gutowski

Psychology Department
Tufts University
Medford, Mass. 02155

and

Everett A. Palmer

Man-Vehicle Research Division
NASA-AMES Research Center
Moffett Field, Cal. 94035

Abstract

An experiment has been done to study the effects on divided attention of visual shifts and long-term memory retrieval during a monitoring task. A concurrent vigilance task was standardized under all experimental conditions. The results show that subjects can perform nearly perfectly on all of the time-shared tasks if long-term memory retrieval is not required for monitoring. With the requirement of memory retrieval, however, there was a large decrease in accuracy for all of the time-shared activities. It was concluded that the attentional demand of long-term memory retrieval is appreciable (even for a well-learned motor sequence), and thus memory retrieval results in a sizable reduction in the capability of subjects to divide their attention. Also, a selected bibliography on the divided attention literature is provided.

*Funds for the support of this study have been allocated by the NASA-AMES Research Center, Moffett Field, California, under Interchange No. NCA2-OR785-801.

INTRODUCTION

This research is an outgrowth of our interest in problems of divided attention that are associated with avionic computer systems. Unlike many other man-computer interactions, the pilot has other time-shared duties besides interacting with the onboard computer. Furthermore, existing avionic computer systems have employed a variety of designs for the man-computer interaction, but virtually nothing is known about the attentional demands of these designs. Some systems are perhaps easier than others to operate in the context of time-shared attention. However, few generalities can be learned from the direct study of existing avionic designs since these designs simultaneously vary on too many dimensions which many affect divided-attention performance. Also, field tests are too unstructured and uncontrolled to learn about the mechanisms affecting performance on time-shared tasks. Thus, in the present research, a controlled laboratory study on divided attention is employed in order to identify the critical dimensions that affect the sharing of attention. Ultimately, the goal of the laboratory studies is the establishment of design guidelines for future man-computer systems. More specifically in this study the number of steps, shifts in visual angle, and information retrieval from long-term memory were investigated.

METHOD

Subjects

Twelve male undergraduates at Tufts University served as subjects. Each subject received $20.00 for their participation in three sessions for a total of three-and-a-half hours.

Design

There were three experimental tasks performed in a time-sharing manner. The experimental tasks consisted of:

1. A monitoring task with three response buttons.
2. A vigilance task with a single response button.
3. A recall task of a four-digit number.

Moreover three independent variables were manipulated corresponding to the parameters of the monitoring and recall tasks. The independent variables consisted of:

1. The physical distance separating the three monitor-response buttons.
2. The sequence required for pressing the monitor-response buttons.
3. The amount of delay between the presentation of the memory item (i.e. the four-digit number) and the recall test. (This delay was scaled in terms of a regularly spaced monitoring stimulus, and henceforth the delay will be referred to as the memory lag. Thus, a memory lag of n has n monitoring stimuli occurring between the memory item presentation and the recall test.)

The parameters of the monitoring task were varied among three experimental sessions. The values for recall lag cut across the experimental sessions and all values were tested at each session.

In session A, the physical arrangement of the monitor buttons was a horizontal line across the top of the console. The buttons were 3 cm apart, center to center, creating a visual angle of approximately 3° for the subject. The sequence required for monitor button presses was free, except that all buttons were to be used equally and no button could be pressed twice consecutively.

Session B differed from A principally in the arrangement of the monitor-response buttons. The configuration of their placement formed a 30 cm equalateral triangle with a button at each corner. The visual angle was approximately 28° between any two buttons.

Session C was also similar to Session A except that a set sequence was required for the monitor button presses. The configuration of the monitor buttons was at the top of the console, as in session A, but the subject was required to press the monitor buttons according to a previously learned pattern.

Apparatus and Procedure

The subject was seated approximately 60 cm in front of a console with 4 buttons and a 9 cm voltmeter face. Three of the buttons were designated as monitor-task buttons and the fourth was designated as the vigilance-task button. All stimulus presentations were under the control of a Sony TC630 stereo tape recorder. The subject's button presses were recorded by event markers on a Narco-Bio Physiograph Six polygraph. All stimulus events were recorded on a separate channel of the polygraph. Recall task stimuli were presented by a Kodak Carousel Projector that was switched by a Lafayette Voice Key. The vigilance task stimulus was switched on by a Uher F422 diapilot. See Figure 1 for a more detailed apparatus schematic.

Each trial began with the 0.5 sec. visual presentation of a random four-digit number. Number stimuli were screened for highly meaningful patterns and those beginning with zero were also excluded. The monitoring task and the vigilance task followed immediately in a time sharing manner. Each trial contained 18 monitor stimuli, presented at a rate of one per 1.25 sec. The monitor stimulus consisted of an audible 0.25 sec. 600 Hz tone which produced an 8 volt deflection at the console meter. The vigilance stimulus was a needle deflection of an additional 6 volts occurring simultaneously with the monitor tone. Six vigilance stimuli were distributed randomly within each trial. The monitor task and the vigilance task were interrupted after lags of 0, 2, 4, 6, or 8 tones for a 3.75 sec. recall test of the memory item for that trial. At the conclusion of the recall interval, monitoring and vigilance resumed for the remaining portion of the trial. The subject was allowed a 5 sec. rest after every trial.

Each experimental session consisted of 90 trials. Eighteen replications of each interrupt lag condition were distributed randomly throughout each session. Each subject participated in all three experimental sessions. The order of participation in session A, B, and C was counterbalanced across the twelve subjects. Prior to the first session, each subject received training on each of the separate experimental tasks and four practice trials in which they were performed in a time-sharing man-

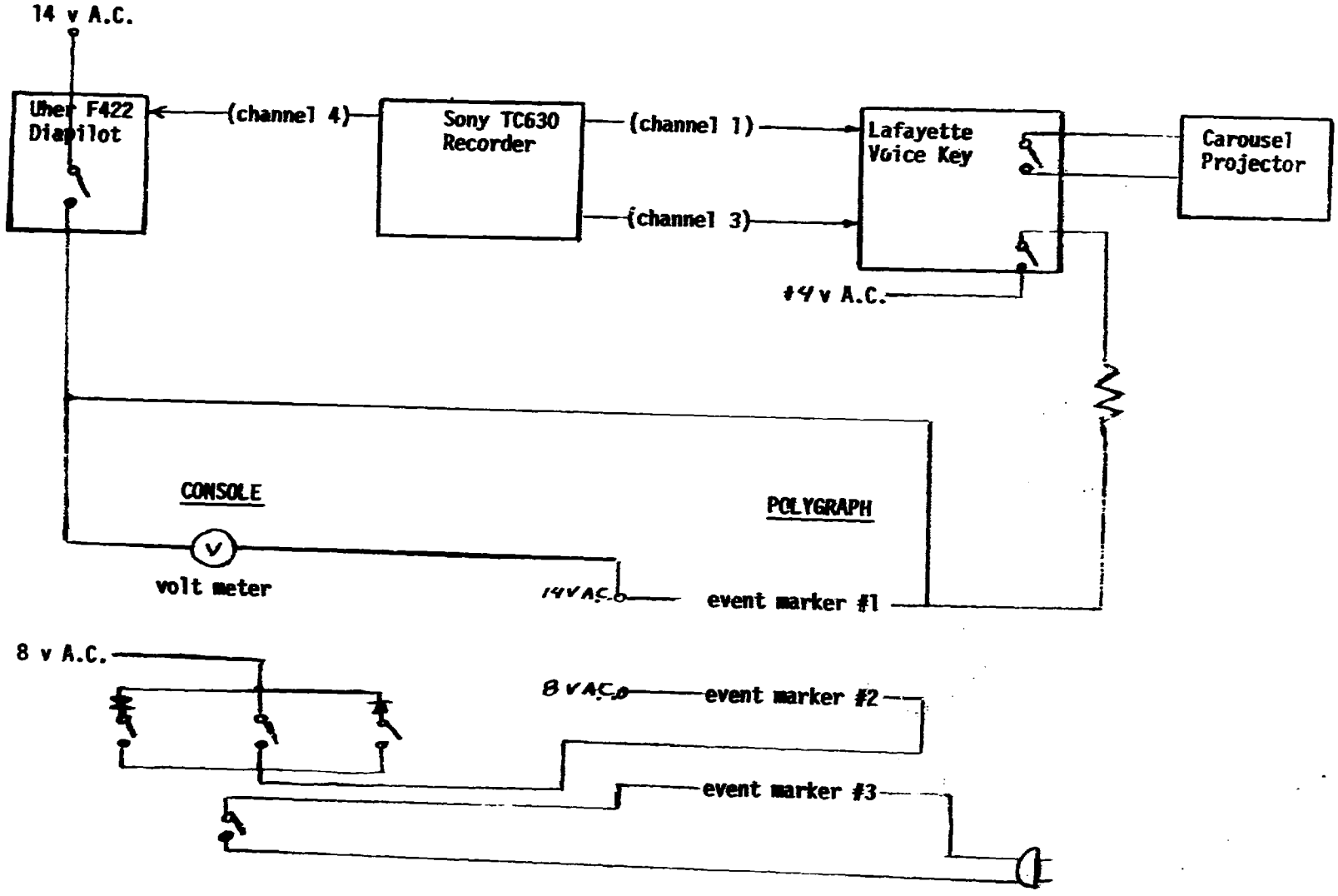


Figure 1 Apparatus schematic

ner. Prior to Session C, the subjects were given one day to memorize the set sequence of 18 monitor-button presses. Also, before Session C began, the subjects had to demonstrate that they, in fact, learned the sequence for monitoring.

RESULTS AND DISCUSSION

The results are broken into three sections corresponding to the three time-shared tasks. Moreover, prior to performing statistical tests, percentage data were transformed by a standard arcsine transformation in order to better assure the conditions of normality.

Monitor Task

The overall results for this task are quite clear as shown in Table 1. Performance is extremely accurate in both Sessions A and B, but decreases considerably in Session C. Since performance was so accurate in Sessions A and B, the data for these experiments were pooled and compared to the Session C data. The mean monitor accuracy is significantly less in Session C, $t(11) = 4.73$, $p < .001$.

TABLE 1

MEAN PROBABILITIES OF CORRECT RESPONSES, INCORRECT RESPONSES, MISSES AND FALSE ALARMS FOR EACH SESSION IN THE MONITOR TASK

	Session		
Response Type	A	B	C
Correct	.996	.994	.794
Incorrect	.002	.001	.175
Miss	.001	.003	.026
False Alarm	.001	.002	.005

In order to explore the nature of the reduction in accuracy in Session C, a number of subsequent analyses were performed. Figure 2 shows the probability of a correct sequence of button presses as a function of sequence length. Note that this curve is necessarily a decreasing function since an error at any monitor step number means that the sequence is also incorrect. Nevertheless, the performance in Sessions A and B is highly accurate for all sequence lengths, but accuracy is lower in Session C for all sequence lengths. The probability of a correct monitor sequence of 18 button presses is greater in Session A and B than the probability of even a one member sequence in Session C. Thus the requirement of recalling the monitoring sequence from long-term memory has a sizable effect on monitoring accuracy.

The actual recall of the four-digit memory item is an interruption from the monitoring and vigilance tasks. Figure 3 shows monitor accuracy relative to the interrupt occurrence. Performance is uniformily high in

Sessions A and B, although there is a slight decrease occurring just prior and subsequent to the interrupt. The data from Session C show that performance is higher before the interrupt than after the interrupt.

In the previous analysis the 0-lag data were omitted since the monitoring and vigilance tasks are not interrupted in the 0-lag condition. Figure 4 displays the monitoring accuracy as a function of monitor step number for the 0-lag condition. This display thus eliminates the effects of interrupts on monitoring. The curve for Sessions A and B is quite flat, again reflecting the high performance level in those conditions; however, the curve for Session C is uniformly lower.

Overall then, there was no effect found on monitoring performance for the spatial arrangement of the monitor buttons. Thus, large visual angle shifts of 28°, that are required in Session B, did not significantly affect the monitoring accuracy. However, requiring the retention of a set monitoring sequence in Session C did produce a dramatic decrease in performance.

Vigilance Task

The overall results for the vigilance task are shown in Table 2. Again, performance is very accurate in Sessions A and B and decreases in Session C. Since performance was equivalent in Sessions A and B, the results of these sessions were again pooled and compared to the data from Session C. The comparison of overall accuracy showed that performance

TABLE 2

MEAN PROBABILITIES OF CORRECT RESPONSES, MISSES AND FALSE ALARMS FOR EACH SESSION IN THE VIGILANCE TASK

	Session		
Response Type	A	B	C
Correct	.997	.995	.974
Miss	.007	.011	.068
False Alarm	.001	.003	.006

decreased significantly in Session C, $t(11) = 7.018$, $p < .00005$. As is evident in Table 2, this performance decrease is almost entirely due to an increase in miss rates since false alarm rates are very low in all three sessions. Further analysis shows that not only the overall rate, but also the pattern of errors changed in Session C. In Sessions A and B 63% of the errors occurred over the first nine monitor step numbers, while only 46% occurred for these monitor numbers in Session C. The memory load in Session A and B is largely provided by only the retention of the four-digit number which is required only for first part of the monitoring sequence. However, for Session C there is a memory load throughout the entire sequence since the monitor sequence itself must be recalled from memory. This difference is consistent with the observed patterns of vigilance errors.

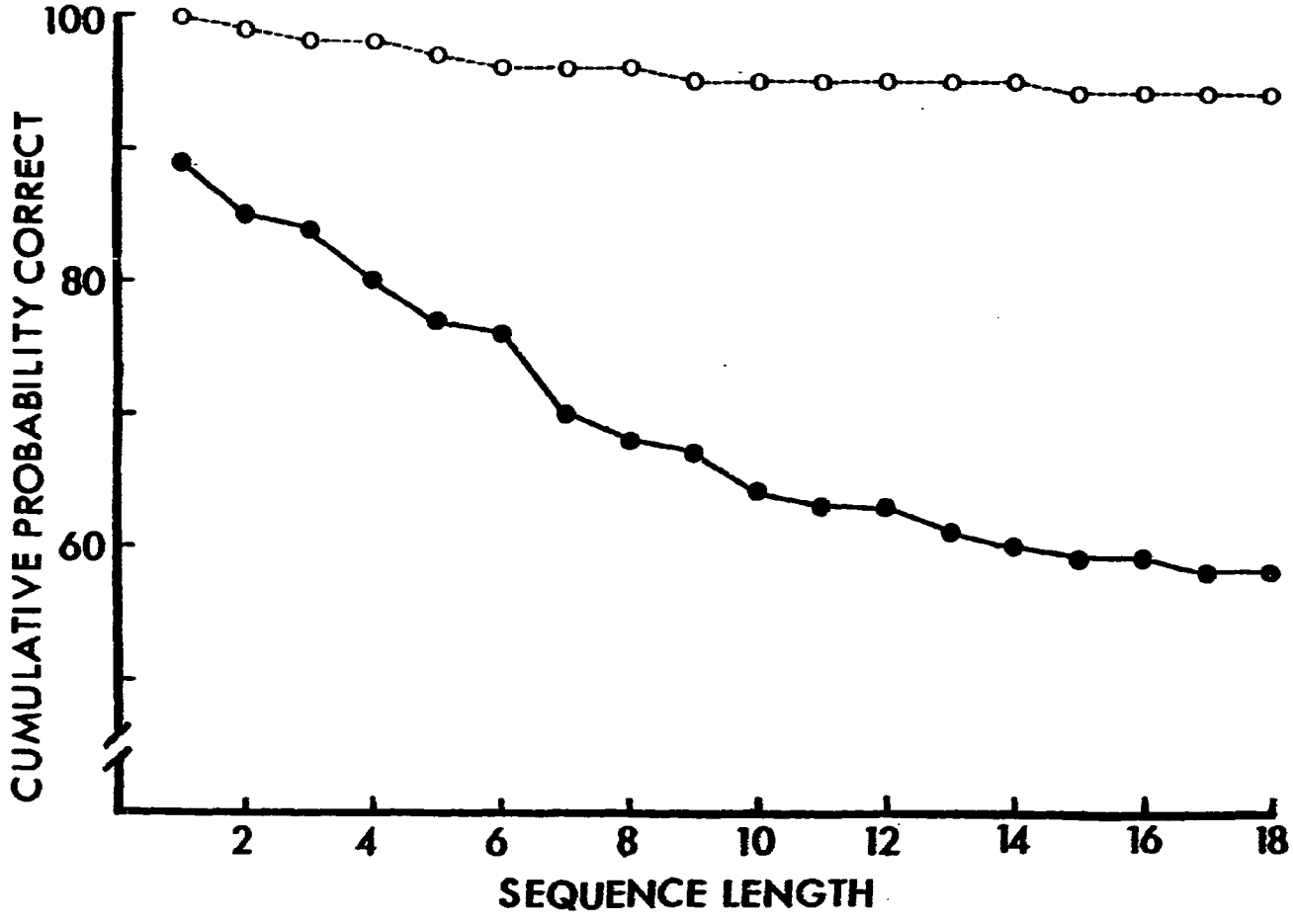


Figure 2. Mean cumulative probability of a sequence of correct monitor responses as a function of sequence length. The open circles represent data averaged over sessions A and B.The closed circles represent data from session C.

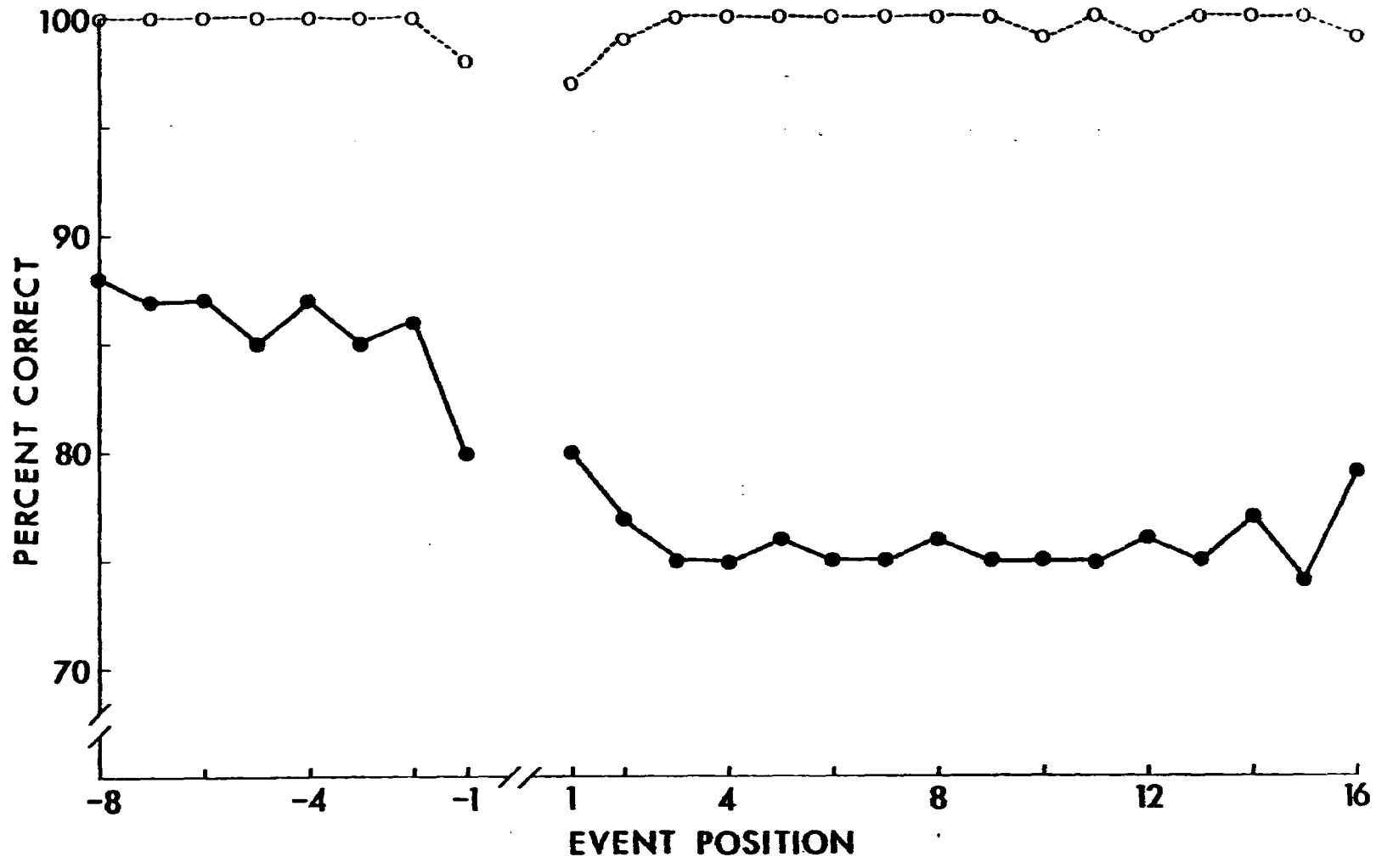


Figure 3. Mean percentage of correct monitor responses as a function of event position. Event position is determined relative to the recall interrupt. The open circles represent data averaged over sessions A and B. The closed circles represent data from session C. Only data from trials with a lag of 2,4,6, or 8 are shown here.

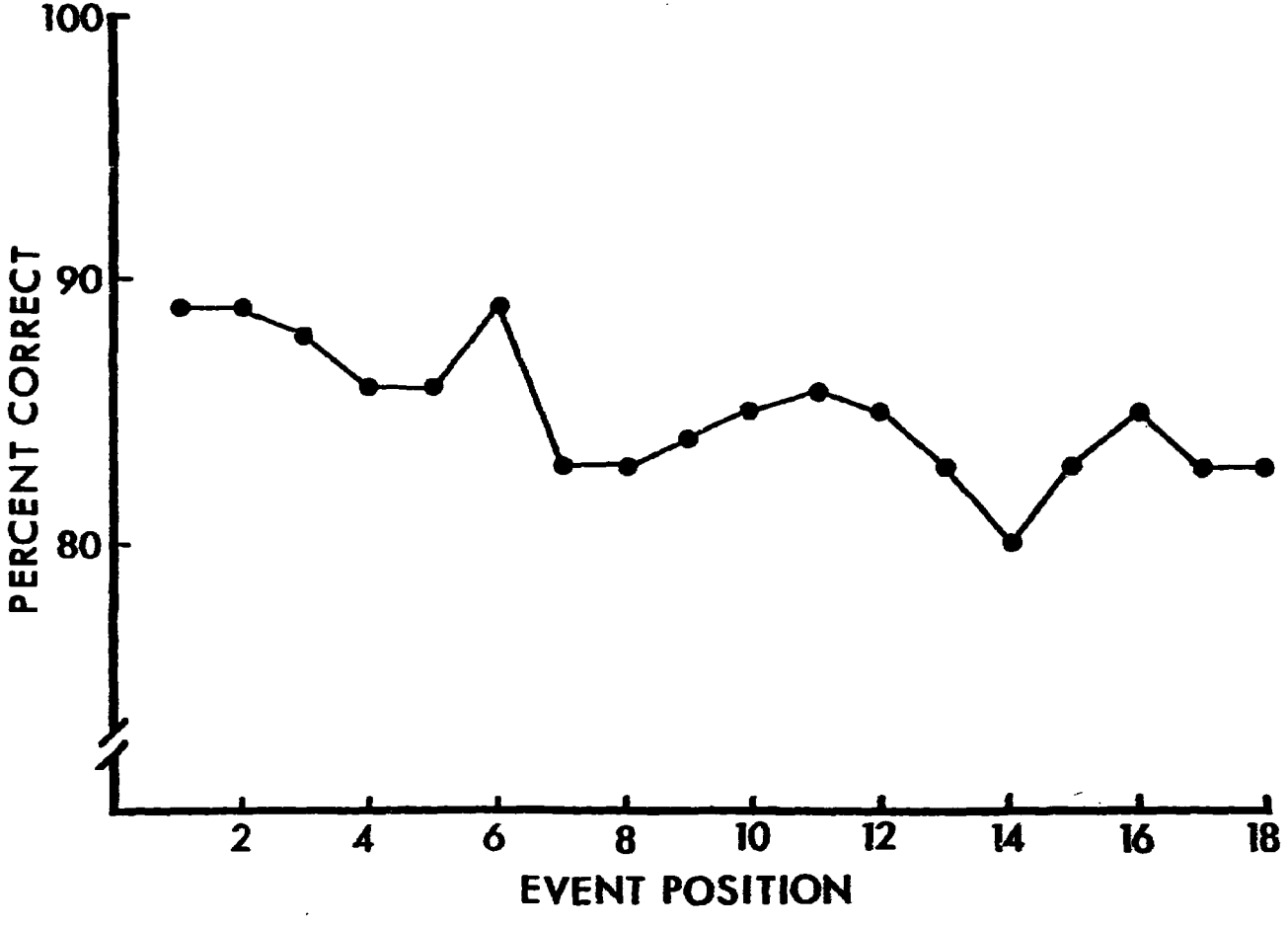


Figure 4. Mean percentage of correct monitor responses on trials with a lag of 0 in session C as a function of event position.

In general the results of the vigilance task are consistent with the monitor results. There is no effect of spatial array found, however, a highly significant decrement in performance occurs in Session C when long-term memory of the monitoring sequence is required.

Results from the Memory Task

The mean percent of recall is shown in Table 3 as a function of session and lag. Analysis of variance was conducted on these data, and it showed significant main effects of both session ($F(2,22) = 8.71$ MSE = .0064, $p < .005$) and lag ($F(4,44) = 6.87$, MSE = .0247, $p < .001$). The session effect resulted from a significant decrease in recall performance in Session C below the very accurate levels maintained in both Sessions A and B. This difference was quite consistent across all lag levels. Although these differences tended to increase as lag increased, there was no significant interaction between session and lag. The main effect of lag is due to the decrease in recall as the number of interpolated events increases. This effect is consistent across all sessions, but somewhat attenuated by ceiling effects in sessions A and B.

TABLE 3

MEAN PROBABILITY OF CORRECT RECALL IN THE MEMORY TASK AS A FUNCTION OF SESSION AND LAG

	Lag				
Session	0	2	4	6	8
A	.991	.986	.990	.977	.977
B	.995	1.00	.990	.990	.976
C	.981	.972	.954	.940	.893

Overall, the results of the memory task are very similar to those found in the monitor and vigilance tasks. Subjects maintain very accurate performance in Sessions A and B, but recall of a set sequence in the monitor task results in a significant decrease in the recall of the four-digit memory item.

SUMMARY AND CONCLUSIONS

The combined results of the monitor, vigilance and memory tasks form a clear pattern. The data from Session A establish that subjects can successfully divide their attention across these three tasks. The results of Session B indicate that increases in visual angle for the monitoring task does not impair performance in any of these tasks. In contrast, requiring subjects to recall a set sequence in the monitor task produces a significant decrease in accuracy in all three tasks. Thus, the information processing of recalling the next action of a memorized sequence of button

presses impairs processing in other concurrent tasks. The retrieval from long-term memory of the next action requires attention that otherwise could have been directed to the time-shared activities. The attentional demand of memory retrieval is appreciable and hence results in a sizable decrease in divided-attention performance. However, without the involvement of long-term memory retrieval, subjects perform nearly perfectly on the time-shared tasks of recalling a recently presented random number and of performing motor responses to monitor and vigilance stimuli.

Attention Allocation in Dynamic Environments

by
Christopher D. Wickens and Pamela Tsang
University of Illinois at Urbana-Champaign
Champaign, Illinois, 61820

ABSTRACT

Three policies of attentional resource allocation between tasks of dynamically varying difficulty are described. These policies--optimal allocation, optimal resource expansion, and non-optimal allocation are distinguished analytically by the gain of the transfer function between task difficulty and primary and secondary task performance. Eight subjects time-shared two compensatory tracking tasks in which the control dynamics of the primary task fluctuated continuously between first and second order. Linear control analysis of the difficulty and filtered RMS error performance measures indicated that subjects were initially non-optimal in their allocation policy, failing to guard the primary task in the face of fluctuations in its difficulty. With practice, a trend toward more optimal performance was observed. However, close analysis and comparison of these data with performance in constant difficulty dual task conditions indicated a persisting limitation in subjects' ability to reallocate resources from the secondary task when required by demand changes of the primary.

INTRODUCTION

When two tasks of similar structure are performed concurrently, it may be assumed that the performance of each relies upon a common pool of processing resources [1,2,3,4,5]. When more resources are allocated to one task, as a consequence of either an increase in its difficulty, or of its required performance level, fewer are available to the concurrent task, and performance of the latter will deteriorate accordingly. The joint representation of concurrent performance of two tasks, as resources are traded off between them is presented in the Performance Operating Characteristic or POC, an example of which is shown in Figure 1a [5]. The vertical and horizontal axes represent performance measures on task A and B respectively, such that good performance corresponds to higher values. Single task performance is represented by the points falling on the axes, while the points within the space correspond to hypothetical performance measures in dual task conditions. Three such conditions are indicated. One in which resources are allocated equally between tasks, one in which the allocation policy favors task A, and one in which it favors task B. The smooth curve connecting the points--the POC--represents the hypothetical frontier of maximum joint performance, across the set of all possible allocation policies between tasks.

Gopher and Navon [1,2] have described how, as the difficulty of one task (task B in Figure 1) is varied, each difficulty level generates a different POC, with tasks of greater difficulty moving the POC closer to the origin. In the specific case of dual axis tracking when the difficulty manipulation employed is the order of the system transfer function (varied from first to second), Gopher and Navon [2] have shown that the set of POC's thus generated form the fan-like shape shown in Figure 1b. The influence of task difficulty on performance of both tasks grows as more resources are allocated to the task of varying demand.

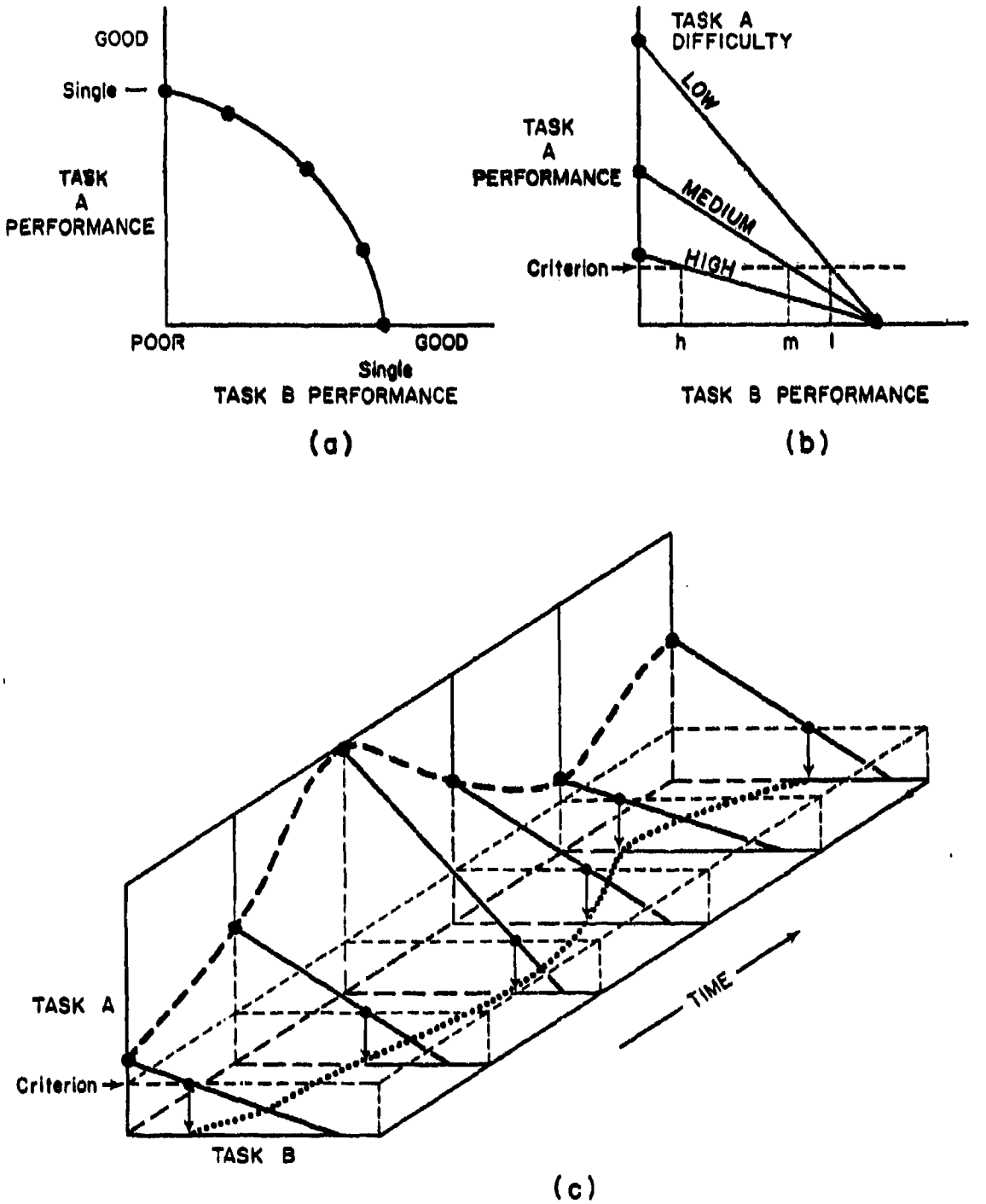


Figure 1. Performance Operating Characteristic in Constant (a), Variable (b), a Time-Varying (c) Difficulty Conditions.

Within the framework of Figure 1b, if task A is designated as primary--to be held at a criterion level of performance (indicated by the horizontal dotted line), then assuming an operator of fixed capacity, this performance can be achieved by trading off resources from task B, the performance of which would be indicated by the points h, m, and l in Figure 1b. Consider now the performance resulting when the difficulty level of the primary task is varied continuously within a trial, rather than discretely between trials. This would be represented by the POC of Figure 1b oscillating between the two extremes of difficulty or, alternatively, in Figure 1c, a time axis is incorporated and the POC now represents an undulating surface. Criterion performance is the horizontal plane that intersects this surface, and optimum secondary task performance, of a fixed capacity system with perfect allocation is represented by the intersection of this surface with the criterion plane projected onto the secondary task (task B) "floor" axis.

An alternative representation of this hypothetical data pattern is shown in the top panel of Figure 2. Wickens and Pierce [6,7] have argued that the transfer function of the inferred resource allocation system can be derived from linear time series analysis of the difficulty and the primary and secondary task performance signals in Figure 2. The pattern shown by the optimum allocator of the top panel would be reflected by a gain, or linear coherence value (between difficulty and performance) that was low for the primary task relative to the secondary. This pattern is referred to as Optimal Allocation. Alternatively, the optimal operator could maintain constant primary task performance by temporarily expanding the supply of available resources at the epochs of peak primary task difficulty. Such expansion has been suggested by Kahneman [8] to be mediated via the role of feedback loops associated with mechanisms of physiological arousal. Such a policy, referred to as Optimal Expansion will, of course, be reflected by reduced values of gain and coherence on both tasks. Finally, a non-optimal response is one in which a fixed supply of resources is maintained to both tasks, and primary task performance varies in coherence with its own difficulty fluctuations. This policy generates primary and secondary task gain values opposite from those of the optimum allocator. In summary, two dimensions of allocation policy may be identified. The degree of optimality is indexed by the difference (or ratio) of the primary and secondary task gain measures, and the degree of expansion by the inverse of the sum of the two gains.

Within this analytical framework, Wickens and Pierce [6,7] observed that the behavior of operators engaged in dual axis tracking fell midway between the categories of optimal and non-optimal allocation. Both gains (and linear coherence) were relatively high and of approximately equal value. In their experiment, the difficulty parameters (the order of the tracking dynamics) were varied in a series of spikes and steep ramps between levels of first and second order. The observation that allocation behavior was far from optimal and did not appear to progress toward optimality across four days of training was surprising and served as the instigation for the present study.

A potential source of the non-optimal response observed by Wickens and Pierce is the severity of the difficulty changes. As a consequence, in the present investigation the difficulty "forcing function" was modified so that pure (non-truncated) sinusoidal components were employed, spanning the range between first and second order dynamics. In addition, the present investigation included a greater number of constant difficulty control conditions than did the prior study, incorporating conditions during which the primary task was maintained at the

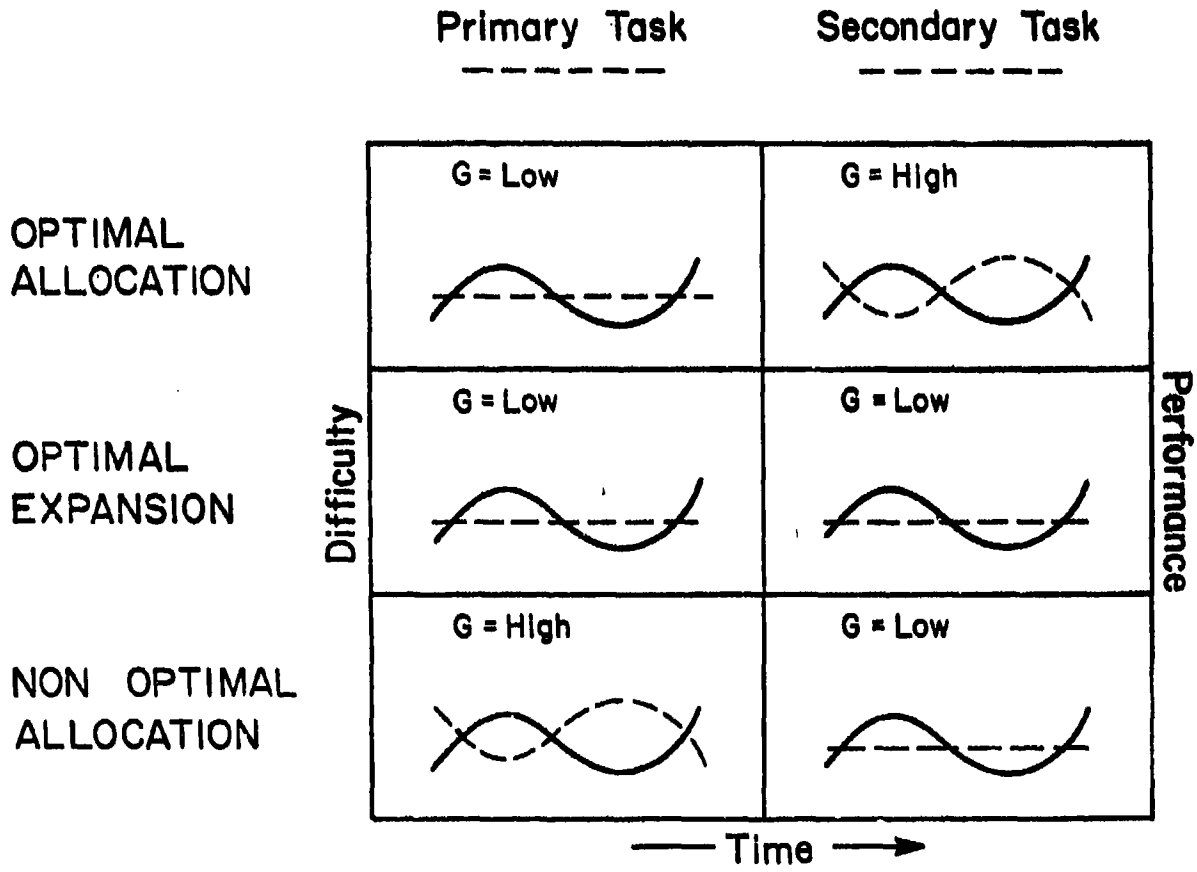

Predicted Gain Values (G) of
Primary Task Difficulty (——)
with performance of the :
Primary Task
Secondary Task
OPTIMAL ALLOCATION
OPTIMAL EXPANSION
NON OPTIMAL ALLOCATION
G = Low
G = High
G = Low
G = Low
G = High
G = Low
Difficulty
Performance
Time

Figure 2

highest, the lowest, and the average level that is obtained under the variable condition.

Method

Subjects performed two compensatory tracking tasks, with dynamics of the form: $Y = K\left(\frac{\alpha}{S} + \frac{\alpha}{S_2}\right)$. The variable difficulty parameter α could thus be set at any value between 0 (first order system) and 1 (pure second order system). Error indicators of the two tasks were displayed laterally on a CRT display with a vertical separation of 0.7 degrees of visual angle. The two displays had a small (1 degree) lateral offset; the right of center display was controlled by left-right deflections of a right hand control stick while the left set display was controlled by a control stick held in the left hand. Both tasks were driven by a separate band limited Gaussion disturbance input with an upper cutoff frequency of .32 Hz.

Six classes of trials employed in the experiment can be dichotomized in terms of whether task difficulty was constant or variable during the trial, and whether single or dual task performance was required. These are represented in Table 1.

TABLE 1

TRIAL TYPES

	Trial Designation	Primary Task Difficulty (α)	Secondary Task Difficulty (α)	
Constant Difficulty	DV	Variable	0.5	Dual Task
	D1	1.0	0.5	
	D.5	0.5	0.5	
	D0	0	0.5	
	S.5	0.5	None	
	SV	Variable	None	

During conditions of variable difficulty, the difficulty parameter α was driven by the forcing function $\alpha = A(\sin F_1 t + \sin F_2 t)$. The value of F_1 was always set at a constant value of .03 Hz. The value of F_2 was varied on alternate trials between .01 and .02 Hz. This variation was incorporated so that subjects would be less able to predict the time course of difficulty variations during a trial. Trials were 200 seconds in duration, allowing two or four full cycles of the lower frequency component. Eight right handed male subjects were employed in a within subjects repeated measures design. On day 1, subjects were provided extensive practice in all task conditions. On days 2-5, following an initial trial of a D.5 type, each subject received the 6 trials in a given order, followed by a 5 minute rest break, and then the same trials in the reverse order. The particular orders used were partially counter balanced across days and subjects.

Instructions emphasized that the right hand task was to be treated as primary. The level of error on that task obtained in the initial D.5 condition was established for each subject on each day as the criterion level. Subjects were

instructed to maintain that level on all subsequent constant difficulty trials, as well as across the deviation of the DV trial. Financial bonuses were employed to reinforce these instructions.

Results

Global RMS Error. RMS error values for the primary and secondary tasks in the six conditions are shown in Figure 3. The figure indicates a decline in error on both tasks with practice, a greater error in dual as opposed to single task performance, and a reduction in this dual task decrement with practice.

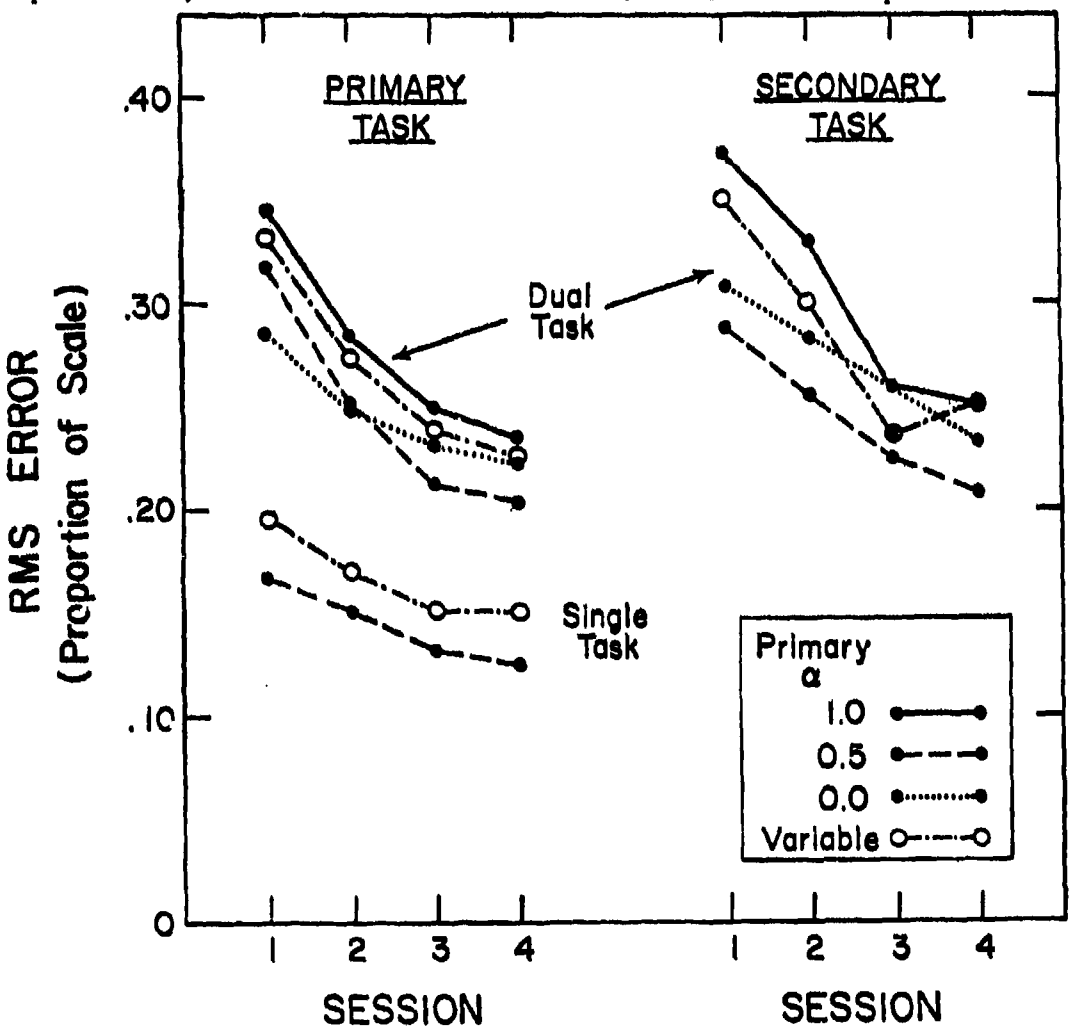


Figure 3

The data were subjected to two ANOVAS. One ANOVA included only the dual task data (both primary and secondary task) and a second included only the primary task data for the single task (SV and S.5) and dual task (DV and D.5) conditions. Both ANOVAS indicated reliable main effects of sessions. The dual task ANOVA indicated reliable effects of α (0, .5, 1 or V) and task (primary vs. secondary) as well,

while in the primary task ANOVA effects of constant vs. variable, dual vs. single, and the dual-single X sessions interactions were all statistically reliable.

Time-series analysis. To evaluate the allocation strategies, the raw sampled RMS error values were smoothed by computing the running average of these values within a 2 second sliding window. The 200 averages computed every second for a given trial were thus used as the performance (output) data in the time-series analysis (Biomed 02T). Separate transfer functions were computed between difficulty (α) and primary and secondary task performance for each subject, in each variable condition on days 1 and 4. Preliminary analyses of the data indicated that the amplitude-ratio measures provided the clearest differentiation between conditions, so only these will be described below. It should be noted, however, that the linear coherence measures obtained here were consistently lower than those observed by Wickens and Pierce [6,7], being of values generally less than 0.50 in the present study.

Figure 4 presents the ensemble average time-series of the primary and secondary task error measures, early (top) and late (bottom) in practice for one of the two difficulty forcing functions employed F=(.02 Hz, .03 Hz). Inspection of

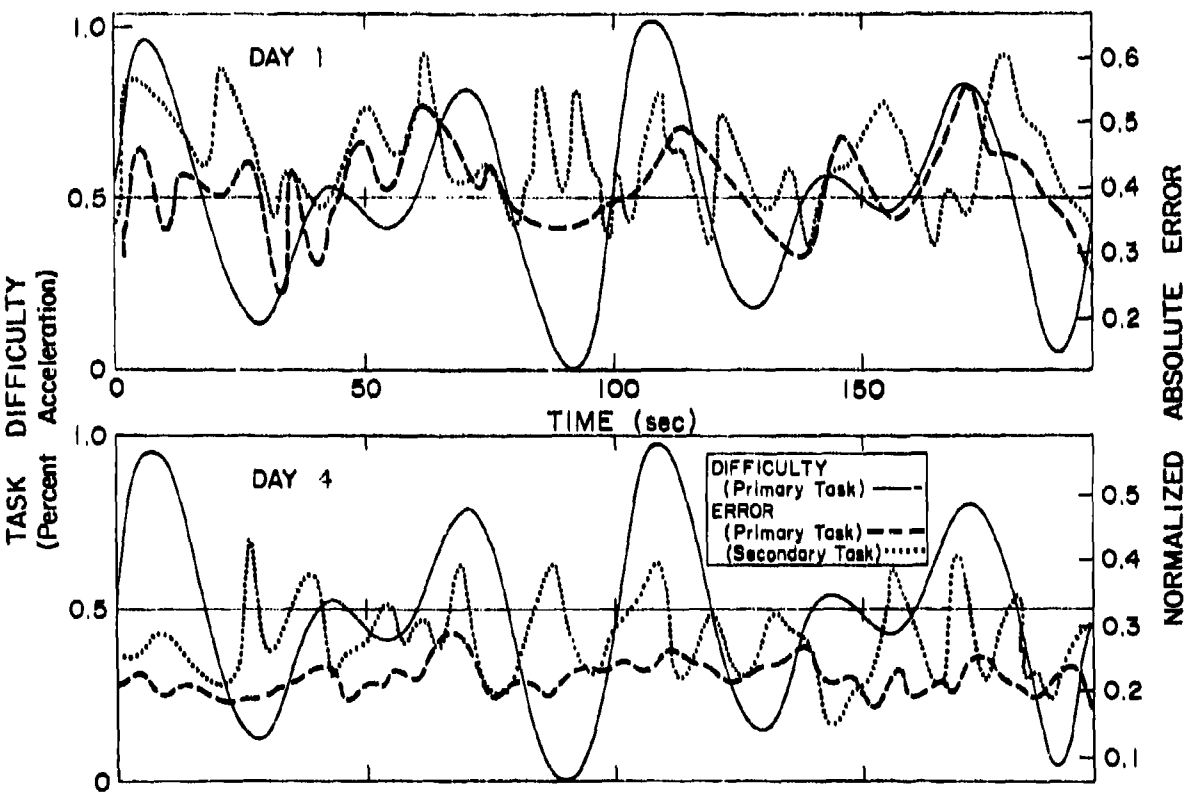


Figure 4. Ensemble Average of Performance and Difficulty Functions.

Figure 4 indicates that the apparent strategy employed early in practice is non-optimal, as primary task performance fluctuates with its own difficulty level. With practice the primary task gain appears to decline; however, this reduction is not paralleled by a corresponding increase in secondary task gain, as would be predicted by adoption of a policy of optimal allocation.

The mean amplitude ratio values of the transfer function at each of the three input frequencies employed across both disturbance functions are shown in Figure 5.

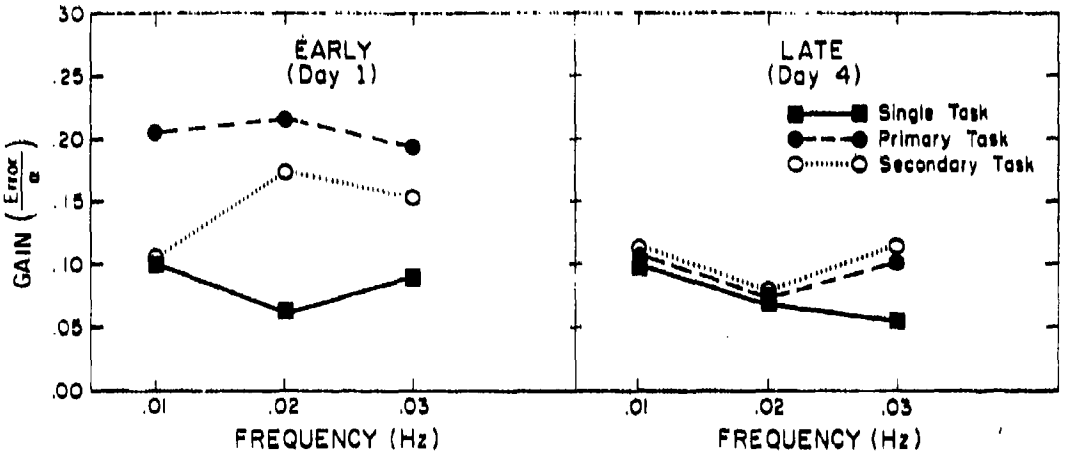


Figure 5. Gain of Difficulty-Performance Transfer Function.

It is apparent that no monotonic trend with frequency is shown in the dual task conditions, suggesting that the allocation system does not behave as a strictly linear system. However, the single, orderly pattern that is observed in Figure 5 occurs with single task gain late in practice. Here there is a suggestion that the response is that of a first order lag, a finding that appears to be consistent with the results of Delp and Crossman [9] in a similar single task condition.

The results of an ANOVA performed on the gain measure indicate that the apparent reduction in gain with practice, evident in Figures 4 and 5, was in fact statistically reliable, as was the greater gain in the dual (primary task only) than single task condition, and the task (dual-single) X sessions interaction. Within the framework of the models presented above, evidence that the subjects were proceeding with practice toward more optimal allocation would be provided by a reliable task X sessions interaction, an indication that secondary task gain increased with practice while primary gain declined. Although this interaction is suggested by the data, its level was not found to be statistically reliable (p = .17). Thus the pattern of behavior demonstrated by the subjects can be described

as one that is initially non-optimal, but manifests a reliable practice trend toward optimal expansion and a non-reliable trend toward optimal allocation.

Constant vs. variable difficulty comparisons. Further analysis focussed upon the apparent trend toward optimal expansion. Operationally, this policy was defined in terms of the ability of the operator to temporarily mobilize more resources than are normally available, during the transient epochs of peak primary task demands. Such a strategy would yield the relatively constant performance on both tasks that was observed. There is, however, an alternative explanation that could be offered for the reduction in both gains. This reduction could result if the resource demands of the difficult ($\alpha = 1$) levels of the primary task were reduced with practice, to a greater extent than those of the easy levels (automation of second order tracking). In terms of Figure 1b and 1c, this explanation posits that the family of POC's draws closer together with practice, thereby lessening the variation in performance on either task. The constant difficulty conditions are crucial for distinguishing these two sources of improvement with practice. If the expansion policy underlies the practice trend and expansion is defined to be transient (thereby evident only in the variable condition), then the "gain" of the performance-difficulty relation should be reduced only in the variable condition. However, if automation is the underlying variable, then a reduction in gain (change in performance with difficulty) that is equivalent in the constant and variable difficulty conditions should be observed.

To enable such a comparison, RMS error values were derived for each subject at the epochs of the DV conditions, when α reached values of 1 and 0, respectively. The difference between these values thereby provides a "gain" measure that has a direct correspondence to the RMSE values in the constant difficulty D1 and D0 conditions. These RMS error values for both the variable and constant conditions, at high and low α, values for the primary and secondary task, early and late in practice are portrayed in Figure 6. The data in the figure suggest that the expansion explanation can be rejected. Late in practice the "gain" in the DV condition is, if anything, greater than in the constant conditions. In a 4-way ANOVA performed only on the session 4 data, this apparent interaction between α and condition (variable-constant) was statistically reliable ($p < .05$). In short, the decrease with practice in the performance difference between the 1 and 0 conditions is just as evident in the constant as in the variable conditions, if not more so. This decrease can therefore be attributed to automation of the task, rather than a change in characteristics of the processing resources themselves.

The source of the larger gain values in the variable, as opposed to the constant difficulty conditions was of some interest. A plausible explanation is that in the constant difficulty conditions subjects were able to adopt a set for each dual task condition, generating the appropriate equalization that is compatable with the system dynamics on that trial and thereby reducing the extent to which performance varies with system order. On the DV trial, however, the lead equalization can be less easily modulated to the appropriate level demanded by the high α periods and so the relatively greater error is obtained.

Conclusions

In conjunction with the previous investigation by Wickens and Pierce (6,7), the present study indicates that human resource allocation in dynamic environments can be far from optimal. If difficulty fluctuations are relatively smooth, as in the present study, then practice does influence the extent to which primary task

performance can be maintained at a constant level. However, comparison of the variable with the constant difficulty control conditions suggested that the mechanism underlying this improvement was not related to an allocation skill, nor apparently to any properties of the resource system itself (e.g., expanded availability). Rather, the reduction in performance sensitivity to difficulty fluctuations seemed to result from an increased automation of primary task performance. This pattern of results casts some doubt on the extent to which resource allocation in dynamic environments may be modelled by a closed loop servomechanism, as Kahneman [8] has argued. Instead, as Galanter [10] has proposed, substantial portions of an operator's response strategies may be generated as ballistic, open loop commands, that are not continuously corrected according to performance feedback. Because of its implications to multi task performance in dynamic environments, the extent to which this apparently non-optimal open loop behavior can be modified by training represents an important area of future research.

REFERENCES

[1] Navon, D., & Gopher, D. On the economy of the human processing system. Technical Report AFOSR-77-1. The Technion, 1977. Also Psychological Review, 1979 (in press).

[2] Gopher, D., & Navon, D. Can you control your control: On the efficiency of attention allocation between tracking dimensions. 15th Annual Conference on Manual Control. Wright State University, March, 1979.

[3] Norman, D., & Bobrow, D. On data limited and resource limited processes. Journal of Cognitive Psychology, 1975, 7, 44-64.

[4] Wickens, C. D. The structure of processing resources. In R. Nickerson and R. Pew (Eds.) Attention and performance VIII. New York: Erlbaum, 1979 (in press).

[5] Wickens, C. D., & Kessel, C. The processing resource demands of failure detection in dynamic systems. University of Illinois Engineering Psychology Laboratory, Technical Report, EPL-79-1/AFOSR-79-1, January, 1979.

[6] Wickens, C. D., & Pierce, B. Linear modelling of attentional resource allocation. Proceedings, 14th Annual Conference on Manual Control. NASA Conference Publication 2060, November, 1978.

[7] Wickens, C. D., & Pierce, B. Attentional resource allocation in a variable difficulty dual task paradigm. University of Illinois, Engineering Psychology Laboratory, Technical Report AFOSR-78-1/EPL-78-1, January, 1978.

[8] Kahneman, D. Attention and Effort, New York: Prentice Hall, 1973.

[9] Dalp, P., & Crossman, E. Transfer characteristics of human response to time-varying plant dynamics. 8th Annual Conference on Manual Control. AFFDL-TR-72-92 June, 1972.

[10] Galanter, E. Human performance capabilities. ONR Final Report. Columbia University, October, 1976. AD A031 462.

Acknowledgements

The research reported was supported by a grant from the Air Force Office of Scientific Research, Life Sciences Directorate Grant AFOSR-77-3380 (Alfred Fregly, Technical Monitor), and a research grant from the National Science Foundation, Cognitive Processes Division No. NSF BNS 78-07860 (Joseph Young, Technical Monitor).

CAN YOU CONTROL YOUR CONTROL: ON THE EFFICIENCY OF ATTENTION ALLOCATION BETWEEN TRACKING DIMENSIONS

Daniel Gopher
David Navon
Technion - Israel Institute of Technology
Haifa, Israel

Abstract

The efficiency of attention allocation between tracking dimensions was investigated in a two dimensional pursuit tracking task where each of the dimensions (horizontal and vertical) was treated as a separate task and manipulated independently. Tracking difficulty on each dimension and their relative emphasis were jointly manuipulated in a central composite response surface design. Negatively accelerated effects of task priority and limited tradeoff between tracking dimensions were obtained when frequency and velocity of target movement served as difficulty parameters. Direct linear tradeoffs were observed when control complexity was increased by changing control dynamics.

These results provide strong evidence to the ability of human subjects to separately control their responses to jointly performed tracking dimensions and allocate varied amounts of processing resources to their performance. The differential effects obtained for the three difficulty manipulations suggest that the locus of load in the time-sharing performance of tracking tasks resides primarily in the response selection stage. It is the complexity of each control response which imposes demands on a common resource and determines the degree of interference between concurrently performed tracking dimensions. Implications of these findings to the design and evaluation of manual control systems are briefly discussed.

Introduction

In multidimensional control tasks such as the control of an aeroplane along the pitch, roll and yaw axes, control difficulty and performance demands often vary dynamically during the flight mission. Pilots are required to adjust their control effort and attend more closely to their performance on the dimension with the increased difficulty. What are the dynamics and empirical consequences of this requirement? Can tracking dimensions be conceived to compete for allocation of a common limited pool of processing and response resources, so that increased emphasis on one axis leads to a commensurate decrease of performance on the other axes? This approach was suggested by Onstott (1976, 1979) in his "Urgency Model" which postuates sequential switchings between tracking dimensions according to priorities established by some urgency function. Onstott supported his model by empirical data collected in the F5E and YF17 flight simulators. Predictions derived from the urgency model disagree with results obtained in other studies (e.g. Wempe and Baty (1969)), in which tracking accuracy on one axis was not impaired when a second tracking axis was added. The amount of interference and performance tradeoffs between concurrently performed tracking axes as a result of joint manipulation of axes priority and tracking difficulty was not tested directly in any of these studies. Thus, several important questions remain unanswered. One question is the extent of the voluntary

control of human operators on the allocation of processing resoucrces among concurrently performed tracking tasks? A related question is the impact of such voluntary changes on tracking accuracy. A third question is the effect of difficulty manipulations on concurrent tracking performance. Increased difficulty on one task increases its demand for processing resources to maintain performance at a fixed level, or performance is impaired if a constant amount is invested and performance is allowed to vary (Navon and Gopher - in press). Comparitive evaluation of the effect of different difficulty parameters on the tradeoff between tracking dimensions appears worthwhile.

These questions were studied in three experiments in which a two dimensional pursuit tracking task was employed as a time-sharing situation. Although only single display and single hand controllers were employed, each of the tracking dimensions (horizontal and vertical) of this task was treated as a separate task and manipulated independently. Both tracking difficulty on each axis and their relative priority were varied. The three experiments differed in the variables employed to manipulate tracking difficulty. In the first experiment, cutoff frequency of target forcing function was employed to manipulate difficulty. In the second experiment velocity of target movements was varied by changing the gain parameters of the forcing functions generating vertical and horizontal velocity vectors. In the last experiment control dynamics was varied to increase task difficulty by incrementing the proportional contribution of second order, acceleration components in mixed first and second order controlled elements. All three are standard manipulations of tracking difficulty (Poulton 1974).

Experiment 1 - Manipulation of Cutoff Frequency

Method

Apparatus

A target and a control symbol were presented on a CRT display (Figure 1). The target was driven in the horizontal and vertical dimensions by two independent, random, band-limited forcing functions (effective screen size was 12 x 12 cm.). The control symbol was controlled through a single two-dimensional, spring centered, hand controller. Control dynamics was a mixture of velocity and acceleration control (for a detailed discussion of this control dynamics, see Gopher, Williges, Williges and Damos, 1975, Wickens and Gopher, 1977)).

Experimental control and data collection were governed by a PDP 11/45 digital computer.

Procedure

Each subject participated in five two-hour experimental sessions. Each session was comprised of 18 three-minute tracking trials and equal periods of rest. The first three sessions were devoted to training and calibration. Adaptive procedures were employed to train subjects under increased levels of tracking difficulty. Adaptive training procedures are computerized automatic algorithms in which trainee's performance is measured and used to set the level of difficulty of the training task. This approach makes it possible to maintain performance of the student at a criterion level as he progresses at his own rate through the range of task difficulty. At the end of this phase, maximum performance measures were obtained for each subject on one and two dimensional tracking.

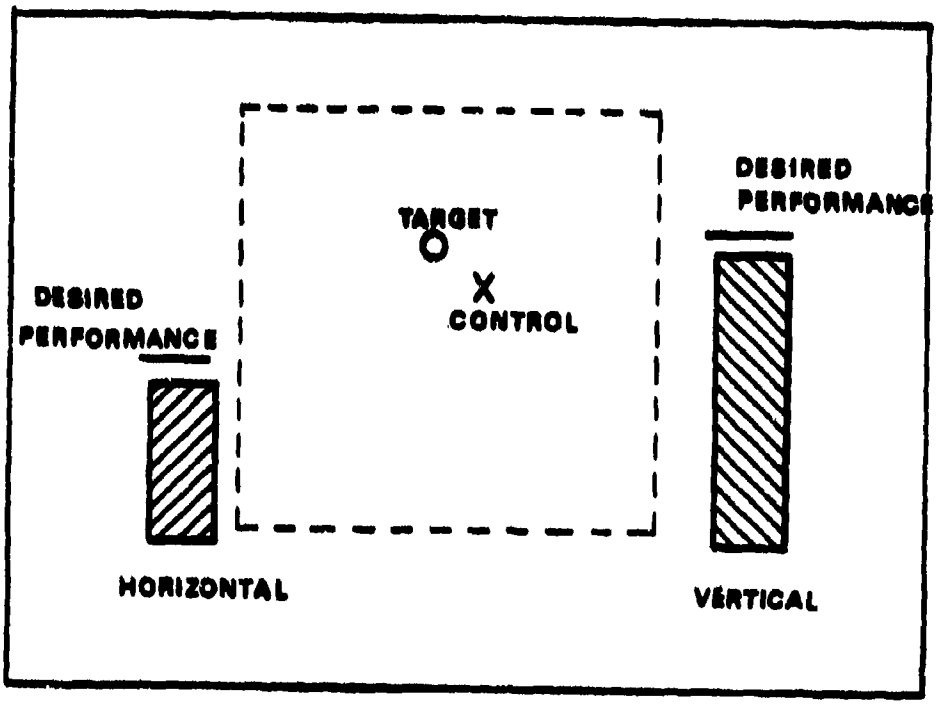


Figure 1 - Subject display with feedback indicators.

The experimented variables were manipulated in sessions four and five.

Design

Three variables were manipulated in the first experiment: One was Task Priorities: These were manipulated by varying the minimal acceptable level of tracking accuracy on each dimension. A second independent variable was the difficulty of horizontal tracking. This was manipulated by changing the cutoff frequency for the low-pass filter applied to the output of a random noise generator to yield the target forcing function. The third variable was the difficulty of vertical tracking, manipulated in an analogous way.

Difficulty Manipulation Tracking difficulty in each dimension could be manipulated manually or adapted continuously by the computer through varying the location of the low-pass, first order digital filters that determined the cutoff freqency of the horizontal and vertical target forcing functions. Increasing the value of the cutoff frequency on a certain dimension caused the target to change directions more frequently and increased the velocity of its movement.

Priorities Manipulations by Feedback Indicators The following is a brief description of the technique developed to present performance feedback and manipulate priorities in dual-task conditions. A more detailed discussion of these techniques can be found in North and Gopher. 1976; Gopher and North, 1977. When in a fixed condition (i.e. forcing functions frequency bandwidth was not adapted), subjects were presented with an on-line, continous feedback on their performance. Feedback indicators (one for each axis) were comprised of a short, static, horizontal line and a moving vertical bar-graph.

The static line represented the desired level of performance (in terms of tracking error), which was determined in reference to a normalized baseline distribution of performance obtained for each subject at the end of the training sessions. Subjects were asked to perform at that level or better. The height of the moving bar-graphs reflected the momentary difference between actual and desired performance. This difference was computed by substracting the momentary error score from the desired score and dividing the outcome by the standard deviation of the baseline distribution. Dimension priorities were manipulated by changing the required level of tracking accuracy on each dimension. Priorities were displayed by means of the height of the desired performance lines. Figure 1 depicts the display in this condition. Indicated are the tracking display, the desired performance lines and the performance bar-graphs. Note that desired performance lines for the two axes are located at different heights, relecting a difference in their relative priorities.

Experimental variables were combined according to a central composite, Response-Surface Design. (Myers, 1971, Williges, 1973). A detailed discussion of Response Surface designs is beyond the scope of this paper. However, due to its relative novelty in behavioral research, its major elements are briefly discussed. Response Surface designs enable under certain assumptions and with proper selection of experimental variables, to obtain a significant reduction in the number of experimental conditions required for a complete factorial design. The main rationale is that with proper selection of levels on each of the experimental factors, it is possible to reconstruct the structure of the variables space with minimal loss to the accuracy of predictions. Naturally, Response Surface designs are most cost

effective in multifactors experiments.

Experimental variables included in a Response Surface design should be quantifiable and measurable such that the only constraints on the selection of experimental levels are those dictated by design considerations. Five levels have to be selected on each variable. Their raw levels are properly transformed to a standard common linear scale to represent the levels +1, -1, 0, +α, -α of this scale. Alpha levels represent the extremes of the range of interest on every experimental variable (say, the highest and lowest values of target movement frequency), and their representative standard score value depends on the number of factors in the experiment, the number of replications of each experimental condition and additional design considerations. For a K factor design, the method requires to sample $2^K + 2^K + 1$ experimental conditions out of the total number of possible combinations, should we conduct a complete factorial design with 5 levels on each factor. (In the present three factors experiment, 15 out of the possible 125 combinations are sampled). Table 1 lists the 15 conditions by their specific combination of standard scores levels. Alpha was assigned a standard value of 1.68. The three variables are Target Horizontal Frequency (Freq. x), Target Vertical Frequency (Freq. y) and Task Priority (Prior).

Table 1. The fifteen experimental conditions (standard score levels, $\alpha = 1.682$).

		Experimental Conditions														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Experi-	Freq.x	1	1	1	1	1	1	1	1	-1.68	0	0	1.68	0	0	0
mental	Freq.y	-1	1	1	-1	1	-1	-1	1	0	-1.68	0	0	1.68	0	0
variables	Prior.*	1	-1	1	-1	-1	1	-1	1	0	0	-1.68	0	0	1.68	0

*Positive values horizontal tracking higher priority.

To define a meaningful range of tracking difficulty on each dimension, it was decided to use for each subject the final levels of tracking difficulty, obtained by him in two-dimensional tracking during training as the center points of the difficulty range (0,0). Lower and upper levels for this range (± α) were determined by adding and subtracting from these values the maximum levels obtained when each dimension was performed singly. The selected range for manipulation of priority was .25 - .75 (the five raw levels on this variable were therefore .25, .35, .50, .65, .75). Two dimensional tracking with equal priorities (.5) served as the center point on this factor.

Desired performance levels (namely the height of the desired performance lines on the display) were determined from the priority values in the following way: a priority level of, say, .75, on a certain dimension, corresponded to a level of performance that assumed the 75 percentile in the baseline distribution of performance of that subject. That is, an instruction to put priority of .75 was actually a requirement to perform at a level better than the lowest 75 percent of the baseline performance levels. Other levels of desired performance were obtained in a similar way. Thus, in both manipulations of difficulty and desired performance, the set of

variable levels employed for different subjects could differ physically but presumably had the same psychological meaning across all subjects. Session four and five constituted two repetitions of the complete Response Surface design. Session four was considered as practice.

Subjects

Five male, right-handed subjects participated in the first experiment. Subjects were paid hourly rates for their participation.

Results and Discussion

Report of experimental results concentrates on the data obtained during the last experimental session. Root Mean Square (RMS) tracking error on each tracking dimension (vertical and horizontal) served as the main performance measure. Tracking errors on each axis were measured every 60 msec. and integrated over 15 sec. intervals. The 12 values obtained in this manner for each 3 minute trial were averaged to yield an overall performance score for that trial. Each of the 3 minute trails represented one of the 15 experimental conditions listed in Table 1.

Consistent with the analysis conventions of Response Surface designs, two standard statistical analyses were conducted on the data occurred for the five subjects in the performance of the 15 experimental conditions. First, a least squares, second order, multiple regression equation was fitted to the data to determine the first and second order effects of the three independent variables on tracking accuracy. Then, an analysis of variance was applied to test the reliability of the obtained regression coefficients. Separate analysis was performed for vertical and horizontal tracking accuracy.

Tracking accuracy on the horizontal axis revealed both first ($F(1,65) = 157.8$; $P<.001$) and second order ($F(1,65) = 31.6$; $P<.001$) effects of task priority on tracking RMS error. A change in tracking difficulty on the horizontal or vertical dimensions did not have reliable first order effects on horizontal accuracy, but small second order effects of both manipulations were observed ($F(1,65) = 3.93$; $P < .06$ for horizontal difficulty; $F(1,65) = 6.20$; $P < .05$ for vertical difficulty). None of the interaction terms between the three experimental variables reached statistical reliability. The joint effect of the three variables on horizontal tracking accuracy yielded a multiple regression coefficient of $R = 0.71$ ($F(9,65) = 21.9$; $P < .001$).

Similar pattern of results was obtained for vertical tracking performance. Task priority had strong first ($F(1,65) = 158.7$; $P < .001$) and second-order ($F(1,65) = 16.48$; $P < .001$) effects. Here again, difficulty manipulations had much smaller effects, although reliable first order effect was obtained for manipulation of target frequency on the vertical axis ($F(1.65) = 4.46$; $P < .05$). The multiple regression score for this dimension yielded a somewhat higher correlation coefficient $R = 0.77$, ($F(9,65) = 22.2$, df 9/65, $P < .001$).

First to be examined in these results is evidence for separation of tracking dimensions. With a single controler and integrated display one can rightly question the ability of human subjects to differentially allocate processing resources to the two tracking dimensions. The strong effect of dimensions priority is the clearest evidence for the time-sharing nature of the situation and the ability of subjects to adjust their response separately on each axis. This outcome is also supported by verbal reports

of the subjects. Other evidence along these lines came from the analysis of the Euclidian error scores ($\sqrt{(RMSX)^2 + (RMSY)^2}$); these were calculated to test the hypothesis that rather than responding separately to each dimension, subjects respond directly to the integrated two dimensional Euclidian distance. Euclidian error scores proved to be a poor measure of subjects' performance. Only 26 per cent of the experimental variance was accounted for by the second order multiple regression equation calculated for this dependent measure, as compared with 51 and 59 per cent that were obtained for the separate measures. It appears safe to conclude that subjects were indeed capable of reacting to the two tracking dimensions as separate, concurrently performed tasks.

To examine the competition between tracking dimensions in joint performance, we have to study more closely the effect of changing axes priority and tracking difficulty.

Effects of task priority were studied by plotting together the second order regression lines of this variable on vertical and horizontal tracking accuracy. Figure 2 shows the effect of changing priorities on each axis separately. The abscissa in this figure represents priorities in favor of the horizontal dimension (PrY = 1-X). The ordinate depicts tracking accuracy (expressed as 1 - RMS error). Each of the two curves was fitted by averaging across the different levels of task difficulty. Actual means for the five priority levels employed in the experiment are depicted by x's and dots.

Three important features can be observed in this figure: first, the effect of priorities is large; its overall range is about 13 per cent of scale RMS error. Second, the effect is quite symmetrical on the two axes. Third, the effect is negatively accelerated. Emphasizing a dimension too much is not necessarily productive.

What is the joint impact of the above effects on overall tracking efficiency in the two dimensional situation? This effect is depicted in figure 3. The sum and difference functions presented in this figure were obtained by adding and subtracting the tracking error scores for the two axes in each of the priority combinations presented in figure 2. Figure 3 depicts a perfectly linear difference function which indicates the faithful responsiveness of subjects to the instruction to change priorities. But, it is also evident that they were much more successful in responding to the lenient tolerance level associated with reduced priorities, than with the stringent requirement when priority was increased. This differential success is clearly revealed in the curvilinear sum function which suggests that joint performance was maximized when both dimensions are attended to about equally well.

Finally, we can examine the effects of manipulating priorities on the concurrent tradeoff between vertical and horizontal tracking by constructing Performance Operating Characteristics (POC) curves, as recommended by Norman and Bobrow (1975). POC curves depict for a fixed level of difficulty of concurrently performed tasks, the results of changing task emphasis on the tradeoff of performance between them. POC's are useful to investigate the degree of competition between task pairs for a common limited resource. In the present experiment, POC's were fitted to the tradeoff between axes by jointly solving the two second-order regression equations (RMS X, RMS Y) for different conditions of priority averaging across all levels of tracking difficulty. Figure 4 depicts the resultant POC.

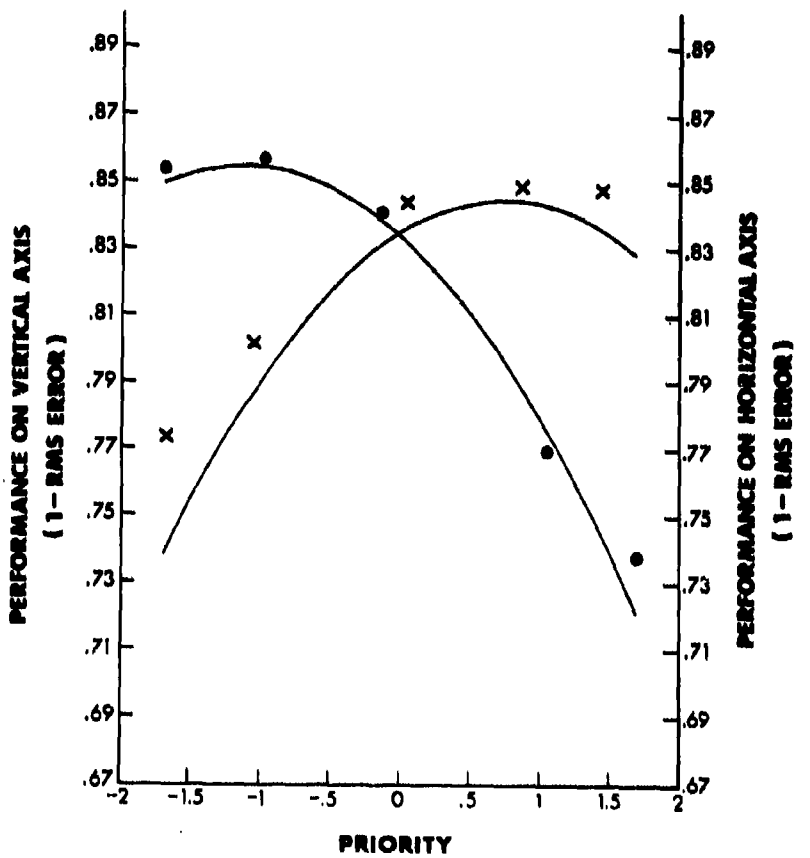


Fig. 2: Effects of task priority on horizontal and vertical tracking accuracy (vertical priorities are expressed as 1-x. Accuracy is presented as 1-RMS Error). Dots represent mean actual data.

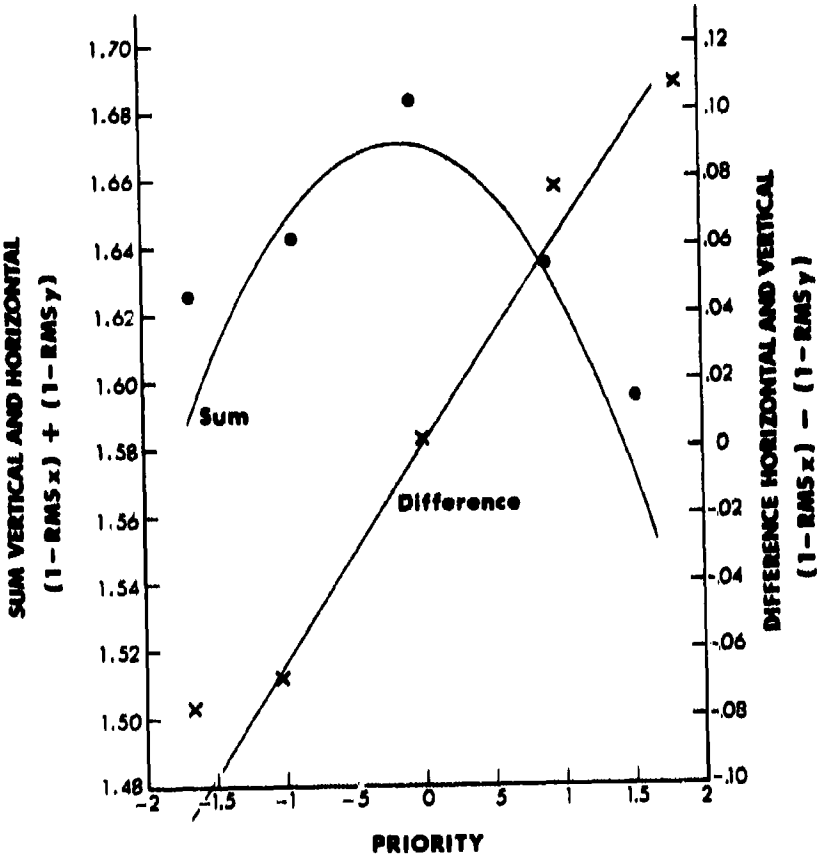


Fig. 3: Sum and difference curves of tracking accuracy scores on the vertical and horizontal dimensions as a function of priorities. Dots represent mean actual data.

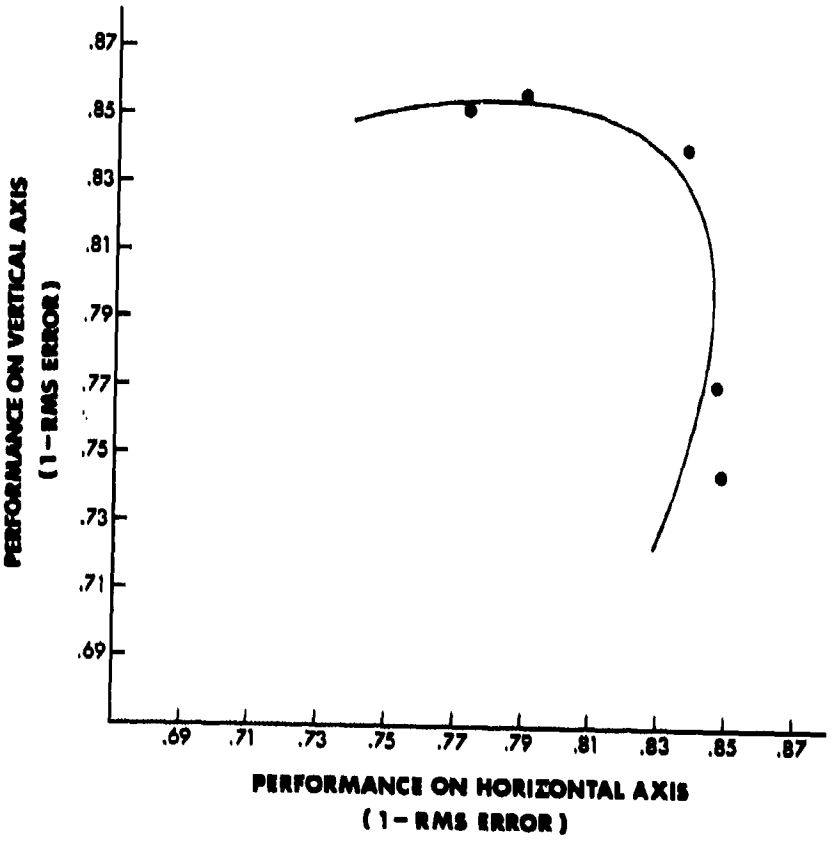


Fig. 4; POC curve depicting performance tradeoff between horizontal and vertical tracking as a function of change in their relative priority. (Average across manipulations of target frequency).

The solid line is the one fitted to the data of all subjects. Dots represent mean actual data for the five priority levels. Performance Operating Characteristics is highly convex, emphasizing the little tradeoff between dimensions, limited to a narrow region of the performance range. The strong curvature of the POC suggests that if task interference is due to a competition of a common resource, performance cannot be linear in amount of invested resources.

An alternative explanation is that the strong curvature and the seemingly lack of tradeoff is due to a ceiling dictated by the nature of the tracking task: subjects simply could not reduce tracking error below a certain level, say .15 - .16 RMS error, no matter how easy the task was, and how much resources were available.

A possible way to test this explanation is to construct a separate POC for different levels of tracking difficulty (replacing the single POC presented in figure 4 with a family of POC's, each depicting the results of manipulating task emphasis for one configuration of tracking difficulty), searching for an interaction between difficulty and priorities. Performance ceiling should be observed in the easy conditions but disappear as task difficulty is increased.

However, interaction between difficulty and priority will be revealed only if tasks share a common limited pool, such that increased difficulty on a certain task would increase its resource demands from this pool for a given level of performance. If the two tasks rely primarily on independent resources for their performance, a change of difficulty on one task would lead to a change of it but would not affect the tradeoffs with the other task. Disjointed resources are therefore a third alternative interpretation of the observed curvilinear relationship between vertical and horizontal tracking. Construction of POC families depicting joint effects of difficulty and priorities are crucial in an effort to distinguish between the three interpretations of the experimental results. In the present experiment, the weak effects of the difficulty manipulation do not encourage such a detailed analysis.

Although reliable, the relative magnitude of the regression coefficients obtained for the vertical and horizontal difficulty manipulations, were about one fourth of those obtained for the priority variable. Different difficulty combinations are therefore not expected to yield a wide spread between curves. Also, because none of the interaction terms between frequency and priorities reached statistical reliability, the shapes and slopes of POC's are not expected to change in different configurations of tracking difficulty and in all conditions should remain highly curvilinear.

It was decided to repeat the experiment with a more potent parameter of task difficulty. Direct increments of target velocity on each tracking axis were employed to manipulate difficulty in the second experiment.

Experiment 2 - Manipulation of Target Velocity

Method

Experiment 2 repeated the design and procedure of the first experiment. The only change was in the manipulation of task difficulty. Target velocity

on each dimension served as difficulty parameter and was varied by changing the gain parameters in the forcing functions controlling target movements. The extreme levels on this manipulation (± α values in the RSM design) were velocity levels representing 30 per cent increase and decrease from the final levels obtained by each subject during the three adaptive training sessions. This procedure resulted in some increase of the general difficulty of the task. Target forcing function frequency on each dimension was fixed at the .7 Hz level. Control dynamics, as in the first experiment, was mixed velocity and acceleration controller. A new group of five male, right-handed students participated in the present experiment.

Results and Discussion

Presentation of results concentrates again on the data obtained in the last experimental session, where task difficulty and axes priorities were jointly manipulated. Separate regression equations and analyses of variance were performed on the RMS error scores obtained for vertical and horizontal tracking.

In this experiment again, and on both axes, the task priority variable had reliable first order effects on tracking accuracy. Second order effects were reliable on the vertical axis and approached reliability for horizontal accuracy. On the horizontal axis: first order effect, $F(1,65) = 41.5$, $P < .001$, second order effect, $F(1,65) = 3.2$; $P < .10$. On the vertical axis: first order effect, $F(1,65) = 36.63$, $P < .001$, second order effect, $F(1,65) = 7.68$, $P < .01$). However, the magnitude of these effects as reflected in the multiple regression coefficient was about half of those obtained for this variable in the first experiment. Table 2 presents the comparative regression coefficients obtained for the priorities variable in the two experiments.

Table 2: Regression Coefficient for Manipulation of Task Priorities in Experiments 1 and 2.

		Regression Coefficients (for standard scores)	
		First ord. effect	Second ord. effect
Exp. 1	Horizontal	2.68	1.80
	Vertical	3.74	1.81
Exp. 2	Horizontal	1.51	0.63
	Vertical	1.30	0.89

Manipulation of tracking difficulty had significant linear (first order) effects on both tracking dimensions (on the horizontal axis, $F(1,65) = 10.43$, $P < .01$, on the vertical axis, $F(1,65) = 28.6$, $P < .001$).

The range of performance change as a result of difficulty manipulations in this experiment was much larger than those observed in the first experiment. Manipulation of target velocity accounted for 7 per cent and 17 per cent of the variance in tracking accuracy on the horizontal and vertical axes respectively. In the first experiment only 2.6, and 1.3

per cent of variance were contributed to the variance in horizontal and vertical tracking accuracy by this manipulation. In addition, no curvilinear effects were revealed and each difficulty variable only affected performance on the axis on which it was applied. As in the first experiment, none of the interaction terms between the three experimental variables was statistically reliable. The multiple correlation coefficients for the effects of the three experimental variables on performance were 0.60 and 0.66 for horizontal and vertical tracking respectively. ($F(9,65) = 6.43$ and 8.37, for horizontal and vertical axes, $P < .001$).

Figure 5 depcits the average POC fitted from the multiple regression equations (averaged across subjects and difficulty manipulations). The POC assumes the same convex shape observed in Experiment 1, reflecting only limited tradeoff between axes. Note the reduced range of performance variability.

With the increased impact of task difficulty, we can now also explore more meaningfully the joint affect of difficulty and priorities. In Figure 6, five POCs are plotted. Each curve represents one difficulty level of vertical tracking, while horizontal difficulty is averaged over all levels employed. All five curves assume the same general shape. Decreasing the level of tracking difficulty on the vertical axis did not change the slopes or shapes of the respective POCs. (Similar results can be plotted for fixed levels of vertical difficulty and different levels of horizontal tracking). POC's asymptote at a lower level for increased levels of vertical difficulty, but vertical distance between curves remain unchanged throughout the whole range. Despite the evident effect of difficulty on subjects' tracking performance, changing the velocity of target movements on one axis did not produce increased tradeoff between axes.

These results seem to overrule a ceiling effect interpretation, because such ceiling should affect only the easier tracking conditions and disappear when task difficulty is increased. They also do not reveal the anticipated interaction between difficulty and resource allocation, if vertical tracking is assumed to compete with horizontal tracking for allocation of a common limited pool. It is as though the two tasks depend mainly on independent resources for their performance. Is there really such a minute competition and interference between tracking dimensions so that performance on both can progress in parallel?

In the first two experiments we manipulated the frequency on velocity of target movement on the screen and control dynamics remained unchanged at a relatively easy level (close to a first order system), in the third experiment control dynamics on the two axes served as the difficulty parameter and the average difficulty of the control task in all tracking conditions was increased.

Experiment 3 - Manipulation of Control Dynamics

Method

Five male, right-handed students served in this experiment. None of these subjects served in the previous experiments. Experimental design and procedures generally repeated those of Experiments 1 and 2. The only

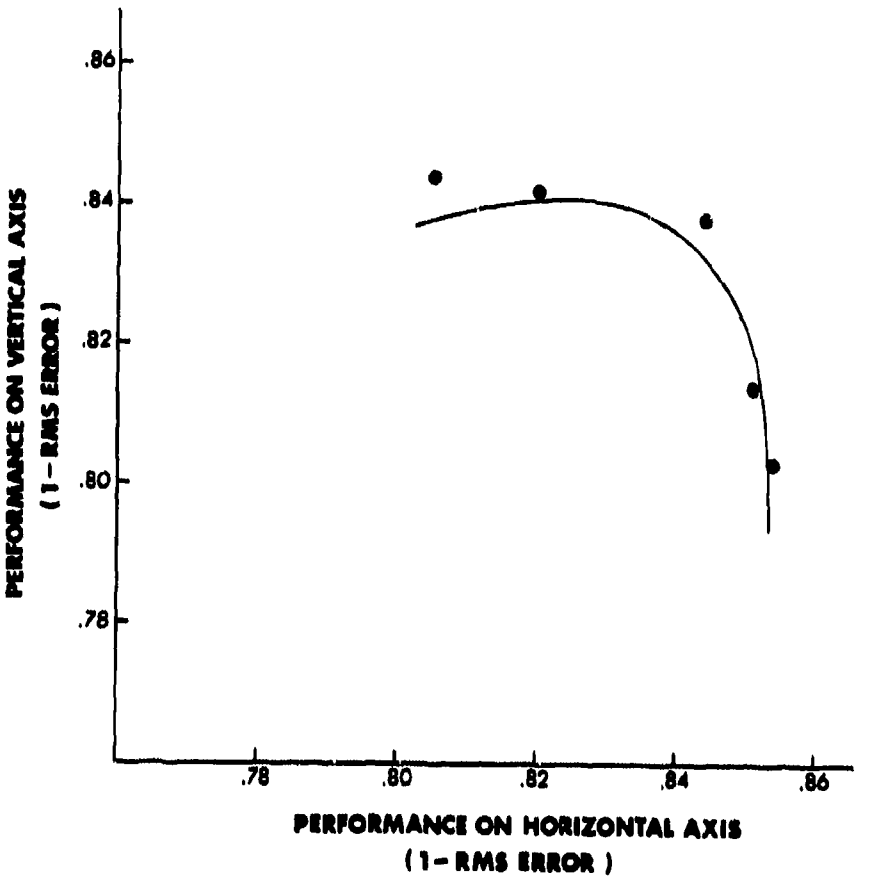


Fig. 5; POC curve depicting performance tradeoff between horizontal and vertical tracking as a function of change in their relative priority. (Averaged across manipulation of target velocity). Dots represent mean actual data.

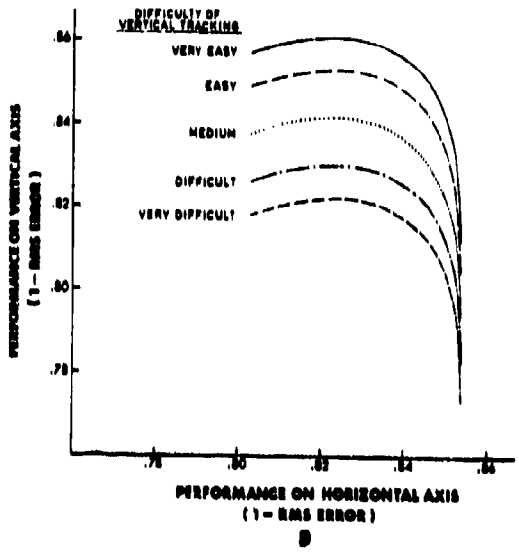


Fig. 6: A family of POCs representing tracking accuracy on each of the axes in dual-axes tracking as a function of task emphasis. Each POC corresponds to a different level of vertical tracking difficulty, (target velocity manipulations). Horizontal tracking difficulty is equal in all curves and represents an average across all difficulty conditions. The five curves were plotted from the regression equations. Because of the sampling strategy of control composite response surface design actual data points could not be systematically added to all curves and were therefore completely deleted.

difference was in the parameter selected for manipulation of tracking difficulty on vertical and horizontal tracking. Control difficulty was systematically manipulated by adding increasing proportions of second order (acceleration) determinants to the transfer functions of the hand controller, thereby incrementing tracking difficulty. Control dynamics generally followed the equation:

$$(1)\ \Theta = (1-\alpha)\frac{.75}{S} + \alpha\frac{.3}{S^2}$$

Theta represents control system output, S and S^2 are the first and second order Laplace transforms. The second order gain (.3) in this transfer function was increased from an 0.2 value which was employed in the first two experiments. Alpha values in equation 1 were manipulated to vary the relative contribution of velocity and acceleration to system response. It is easy to observe that when α equals 0, the second term disappears and the systems react as a pure first order velocity controller. When α equals 1 the system becomes a pure second order acceleration controller, an unstable system which is very difficult to control. (Because an increment in the level of α increases the effect of the acceleration component, the task difficulty parameter in this experiment has been named Acceleration).

During the three preliminary training sessions subjects practiced this variable adaptively until they reached their maximal level on single and dual axis tracking. As in the second experiment the manipulation range on this variable was defined as 30 per cent increase and decrease of the final level reached by each subject at the end of training. This procedure brought again an average increase in the general difficulty of tracking as compared with tracking in the first experiment.

Results and Discussion

Among the three difficulty manipulations investigated in the present study, changing control dynamics appears to create for subjects the most demanding configuration to cope with. This increased difficulty was reflected in an increased instability and noise in the measurement of performance and larger effects of individual variability. As a result, multiple correlation coefficients for the regression equations computed on vertical and horizontal tracking scores decreased to 0.48 and 0.57 respectively. Still, the analysis of variance showed high reliability for both coefficients. (For horizontal tracking, $F(9,65) = 6.01$, $P < .001$. For vertical tracking, $F(9,65) = 8.45$, $P < .001$). Task priority had reliable first order effect on horizontal tracking accuracy ($F(1,65) = 17.18$, $P < .001$), but no second order effect. On the vertical dimension first order effect was highly reliable ($F(1,65) = 50.05$, $P < .001$), and second order effect approached statistical reliability ($F(1,65) = 3.92$, $P = .06$). Clearly, the main effect of priority on tracking performance in both axes was linear and the curvilinearity contributed by second order components in the first two experiments almost disappeared. The overall range of performance variation contributed by the priority manipulation resembled the range observed in the second experiment and was much narrower than the range obtained in the first experiment. The standard regression coefficients for the first order effect on horizontal and vertical tracking were 0.93 and 1.14 respectively. Control dynamics turned out to be a strong and effective manipulation of task difficulty. Highly reliable first order effects were observed for this manipulation on both axes. (For horizontal

tracking: $F(1,65) = 29.69$, $P < .001$. For vertical tracking: $F = (1,65)\ 17.51$, $P < .001$). No second order effects were found.

In addition to the direct effects of difficulty and priority, in this experiment a noticeable effect of interaction between resource allocation (priority) and task difficulty seemed to emerge on the vertical dimension, and approached the common level of statistical reliability ($F(1,65) = 3.23$, $P < .09$).

Figures 7 and 8 present the average POC and a family of POCs plotted for these results. It is evident from looking at these figures that within the relatively shrunken performance range presented in Figure 7, tradeoff between the two tracking axes is close to linear. The interaction between difficulty and priority is manifested in the fan-like family of POCs presented in Figure 8. As tracking difficulty on the vertical axis is decreased, the slope of the POC curve increases. It is in this last experiment that the prediction of interaction between task difficulty and priority as a result of sharing a common resource was borne out.

Implications and General Discussion

To summarize the main findings of the three experiments: subjects were shown to be able to respond separately to each of the dimensions of a two dimensional tracking task, even though integrated display and single controller were employed. In the first two experiments, where frequency and velocity of target forcing functions were varied to manipulate task difficulty tracking performance on the horizontal and vertical axes, did not compete with each other. Subjects could relax their performance one one axis but were unable to use the released resources to improve performance on the other axis on which demand was simultaneously increased. Direct tradeoff between axes was observed in the third experiment when control difficulty was increased and required the subjects to employ a more complex control strategy.

From a theoretical viewpoint, these results suggest that the locus of interference in the time-sharing performance of tracking tasks, resides primarily in the response selection stage. It is the complexity of control response and control strategy on each dimension which seems to impose demands on a common resource and determine the degree of interference between concurrently performed tracking tasks. To account for these differential effects, a refined approach to the tracking task is required. Suppose that, despite the apparent similarity between vertical and horizontal tracking, they do not call exactly for the same types of resources. Furthermore, suppose that the two tasks require the same kind of motor-related resource, but different kinds of perceptual or "computational" resources. In the first two experiments, the load on the motor system (the common resource) was relatively small, because control dynamics was primarily first-order and easy to handle, hence the tasks did not interfere very much with each other. However, with the increased complexity of the control dynamics in the third experiment, both tracking dimensions required more motor capacity in all experimental conditions, so that they had more to compete for. The discrepancy between the data of the three experiments may be understood if we realize that the

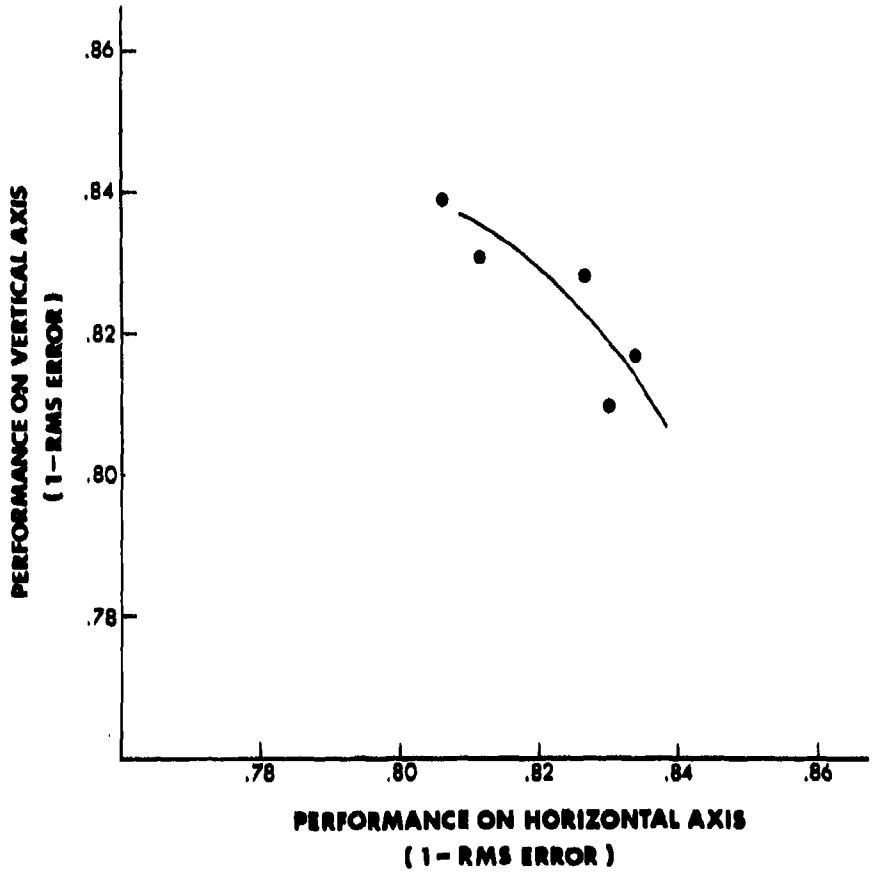


Fig. 7: POC curve depicting performance tradeoff between horizontal and vertical tracking as a function of change he relative priority. (Average across manipulations of control d s). Dots represent mean actual data.

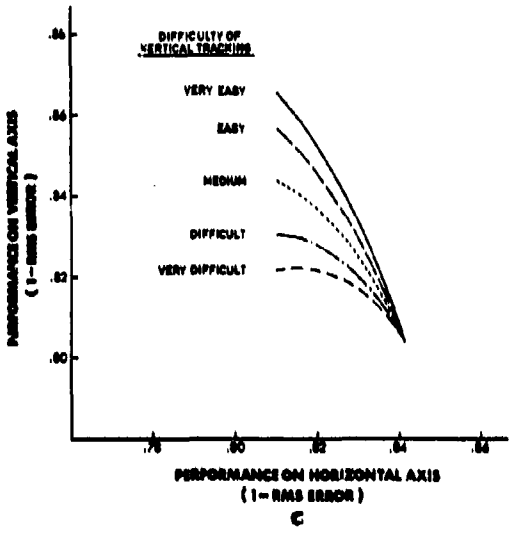


Fig. 8: A family of POCs representing tracking accuracy on each of the axes in dual-axes tracking as a function of task emphasis. Each POC corresponds to a different level of vertical tracking difficulty. (Control dynamics manipulations). Horizontal tracking difficulty is equal in all curves and represents an average across all difficulty conditions. The five curves were plotted from the regression equations. Because of the sampling strategy of central composite response surface designs actual data points for all curves could not be systematically added and were therefore completely deleted.

only parameter which seems to affect the motor system is the control dynamics manipulation. Manipulations of target frequency and velocity of movement did not increase the general difficulty of system control and therefore did not affect the resource utilization of the common resoucre, and had little impact on the tradeoff betwen vertical and horizontal tracking.

The results of the present experiments have several practical implications. They suggest that as long as the control resource is not exhausted, performance on two concurrently performed tracking dimensions can progress in parallel and each of them has separate limitations.

In the evaluation and design of manual control systems, one should be very cautious in the selection of performance measures and determination of load imposed on the operator. Vector error measures are clearly inadequate when attention demands can change dynamically. Employing an integrated control and display device does not necessarily imply task integration with regard to the performance strategy selected by the operator. It also does not allow the designer to assume that every change in the system design would affect the control task as a whole.

These distinctions are of special importance in the design of automatic control augmentation and development of training devices.

References

Gopher, D., North, R.A. Manipulating the conditions of training in time-sharing performance. Human Factors, 1977, 19, 583-393.

Gopher, D., Williges, B.H., Williges, R.C., Damos, D.L. Varying the type and number of adaptive variables in continuous tracking. Journal of Motor Behavior, 1975, 7, 159-170.

Myers, R.H. Response Surface Methodology, Boston: Allyn & Bacon, 1971.

Navon, D., Gopher, D. On the economy of the human processing system. Psychological Review (in press).

Norman, D.A., Bobrow, D.J. On the analysis of performance operating characteristics, Psychological Review, 1976, 83, 508-519.

North, R.A., Gopher, D. Measures of attention as predictors of flight performance, Human Factors, 1976, 18, 1-14.

Onstott, E.D. Task interference in multiaxis aircraft stabilization Proceedings 12th Annual Conference on manual control, NASA TM-X-7, 170, May 1976.

Onstott, E.D., Faulkner, W.H. Definition and Interpretation of flying Qualities, Northrop, TR, NOR. 79-22, March, 1979.

Wempe, T.E., Baty, D.L. Human information processing rates during certain multiaxis tracking tasks with a concurrent auditory task. Fourth Annual NASA University Conference on Manual Control NASA Scientific and technical information division 1969.

Wickens, C.D., Gopher, D. Control theory measures as indices of attention allocation strategies. Human Factors, 1977, 19, 349-365.

Williges, R.C. (Ed.). Special issue: Response surface methodology. Human Factors, 1973, 15, No. 4.

Acknowledgements

The research presented in this paper was supported by the Life Sciences Program, Air Force Office of Scientific Research, under grant No. AFOSR 78-3131. Major Jack Thorpe in the Life Sciences Directorate and Captain Robert Powell in AFOSR European Office were the scientific monitors of this grant.

A DISCRETE CONTROL ANALYSIS OF COORDINATION ACTIVITIES IN A SIMULATED AAA SYSTEM

by
R. A. Miller
Department of Industrial & Systems Engineering
The Ohio State University
Columbus, Ohio 43210

ABSTRACT

The discrete control modelling methods used to represent and analyze performance data obtained from a three person AAA tracking simulation are discussed. The basic model structure, a hierarchical network of communicating finite systems, is presented and selected state transition function estimates are also presented. From the analysis several generalizations about the strategies employed by the subject teams are identified.

1. INTRODUCTION

The general class of systems of interest is that in which the operator (operators) of the system has (have) available only a finite number of control or decision alternatives and these are used to directly or indirectly control the behavior of the system over time. An operator might also have other tasks including continuous control tasks to perform, but the issue here is the set of discrete decisions by which the system configuration and mode of operation is established, and the procedures by which the team members' activities are coordinated.

The modelling questions focus on the problems of capturing in some mathematical representation the way in which team members might decompose a complex problem into simpler parts and how they then manage to coordinate their individual activities and configure the system so that acceptable overall system performance is achieved. The basic questions therefore are questions of knowledge representation, information flow and communication in a complex system. A general hierarchical/heterarchical structure which allows for structural coordination of subsystems by upper level components and which utilizes a heterarchical control structure to shift the focus of control to the proper subsystem at the proper time has been developed. This general structure was used to guide the analysis of a simulated anti-aircraft artillery system. A specific realization of the structure was constructed and the data obtained

from experiments using the above mentioned simulator were interpreted using this structure and specially designed analysis routines. The model and its key properties form the major part of this paper.

2. A DESCRIPTION OF DISCRETE CONTROL MODELLING

The only property established by the definition of discrete control cited above is that a finite number of control alternatives are available for use by the operators. The operators presumably change the alternative (control) selected from time to time in response to changing requirements or a changing environment. Basically, the purpose behind constructing a discrete control model is to explain how specific selections are reached from information about the system, environment and the context of the situation.

It is assumed throughout this development that the information about the controlled system and the environment which is displayed to the operator, or otherwise provided him, is discrete. The assumption is that such data are naturally discrete or are used by him in discrete form. It is assumed that any continuous information can be categorized in some way; e.g., slow, medium, fast. This process of representing continuous information in a discrete qualitative format in some sense corresponds to a feature extraction or abstraction process performed by an operator when encoding and internalizing the information provided him.

Abstractly then the discrete control context consists of an operator or team of operators receiving information in the form of event sequences and producing some sequence of control selections in response. A behavioral representation of a discrete controller (operator or operators) is then a system S,

$$S \subseteq A^T \times B^T$$

where T denotes the time set, A the input alphabet (set of possible inputs to the operators, here assumed finite), and B the output alphabet (set of decision alternatives, also assumed finite). The set A^T is the set of possible input functions with domain T and co-domain A. B^T is similiarly defined. The system S therefore is the relation consisting of the pairs of input-output behaviors which are possible. This relation can be thought of as the sample space from which all data items in a discrete control experiment are selected.

The task in modelling is to provide a more detailed and constructive description of the relation S. In systems theoretic terms this is usually accomplished through a state decomposition of the system and the construction of state transition and output assignment functions. Since the input and output alphabets are both finite sets in

discrete control situtations, an event based state description is most useful. Let C denote the state space, a finite set, and define

$$M = \{m \mid m : C \rightarrow [0,1], \sum_{C} m(c) = 1\}.$$

M should be interpreted as the set of all state occupancy probability distributions defined on state space C. Now a function of the form

$$\underset{\sim}{\Phi} : C \times T \times A \times T \rightarrow M$$

is a state transition function of a very special kind. Specifically,

$$\underset{\sim}{\Phi}(c, t, a, s) \quad M$$

defines a conditional state occupancy distribution with the conditioning provided by the following information:

1) The state at the last event occurrence was c.
2) State c had been occupied t units at that time.
3) The next state change event occurs within s time units.
4) The input over the interval from the last event to the next event is fixed at a.

The function $\underset{\sim}{\Phi}$ then establishes the probability of state occupancy within a specified length of time given certain state and input information.

An output assignment function is simply a function of the form

$$\lambda : C \rightarrow B$$

and it simply assigns the appropriate output symbol to each state. Together the state transition and output assignment functions provide a stochastic representation of the system S. More precisely, when modelling one must find a state space C and functions $\underset{\sim}{\Phi}$ and λ which provide an adequate representation of a given system S.

The concept of state and state transition used above is a highly abstract one and for any complex system requires some additional refinement. Experience with discrete control modelling has shown that a very powerful way of constructing representations is through the use of networks of simple systems.

The diagram shown in Figure 1 is a simple illustration. Each system in the network is a finite state system represented by the usual objects: an input alphabet, output alphabet, state space, transition function and output assignment function. In other words, each system in the network inherits all of the properties discussed above.

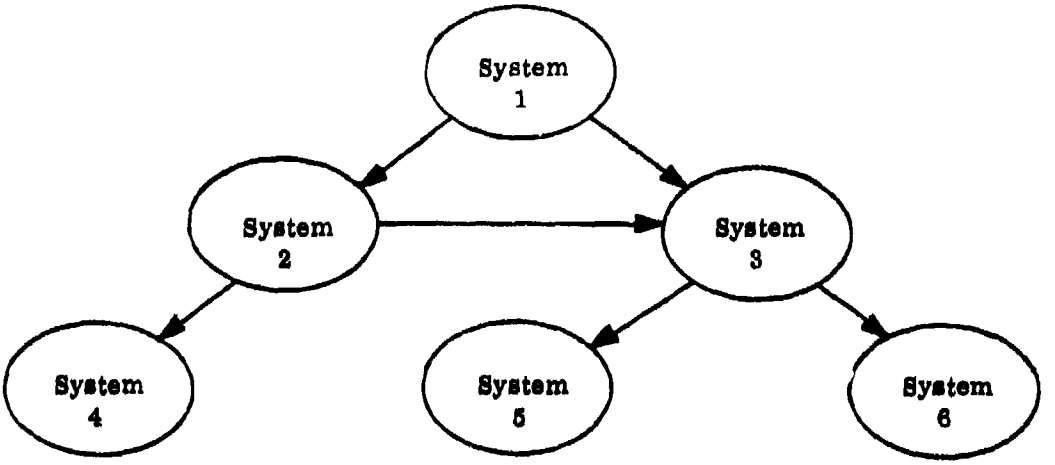


Figure 1.--Network of Systems.

The state space of the complete system; i.e., the network, is then the cartesian product of the state spaces of the component systems. The state transition function and output assignment function required to represent S also follows from the functions associated with the individual elements in the network. Each system in the network is governed by a transition function and any change in state is communicated to the appropriate systems via the communication links defined by the network. The system transition function then is really a fairly complex function which is in some sense the product of the appropriate individual function. But, the important point is that each individual element is quite simple and the rules for constructing the overall state transition function are straightforward. In other words, by knowing the properties of the component systems and the road map which defines their interconnection, the system properties easily follow.

One advantage of the network construction is complexity reduction. The number of transition probabilities that must be estimated is substantially smaller than would be the case without the network decomposition. The network representation also has a distinct substantive advantage. The primary reason for constructing the model in the first place is to help explain how operators perform discrete control tasks. The network is in essence a representation of the intelligence which might be brought to bear on the problem. It is one way of representing a complex problem in a manageable form. Each system in the network represents some important activity or subsystem which the operator must control. By identifying what these component systems are and how they interrelate, the discrete control model is constructed, and network representation follows as a natural by-product.

The main effort in discrete control modelling is spent in constructing the network. Once the network is obtained data analysis and the estimation of transition functions can proceed in a fairly mechanical way, but the analysis must start with the network and the success or failure of the modelling effort depends to some degree on the care which goes into the specification of the structure. A few comments about the overall process of discrete control modelling are provided below.

The first step is to determine all of the discrete outputs which the system is required to specify. These normally are the specific decision alternatives which the operators can select from and typically include items like switch settings and other discrete status indicators. Such items generally can be obtained from a detailed analysis of the system which the operators control. In some cases it may not be necessary or desirable to work at the level of individual switches in which case the analyst must define the proper level and specify in unambiguous terms exactly what the output primitives are to be. The individual items identified in this phase of the analysis determine the system output alphabet.

The second step is to identify the exogenous input variables which in some sense drive the system. These might include things like target trajectories, or command information from other systems. Some of this information will probably be in the form of continuous variables in which case rules for interpreting such data in events format must be defined. This step corresponds to some type of feature extraction through which the essential information classes are extracted from the data. For example, targets might be classed as maneuvering or non-maneuvering as a function of their time behavior. In essence the task is to abstract out a small number of information classes which can then be used for discrete control analysis.

At the same time other non-exogenous continuous information such as tracking errors must be represented in events format. There are no preset procedures for accomplishing this but rules specifically designed to match the problem context must be defined.

The next step requires that the elements to be used in the network be defined. It is important to note that, for purposes of data analysis, the state of any system defined must be computable from available information. That is, data analysis cannot proceed if the state of one or more systems in the network cannot be uniquely specified. With this constraint in mind, the process of defining the required or desired systems proceeds in several stages which are often patterned after a level of abstraction hierarchy. First level systems (components) are one level of abstraction away from the primitive data items and they consist of fairly independent subsystems. These can be established on functional grounds or for purposes of forming aggregate information about the primitives. Second level systems are formed in a similar manner from the primitives and the first level systems. These can be formed to provide coordination of the lower level activities and functions, or they can again simply be an aggregation. This procedure of subsystem definition continues until no further systems are needed.

The key point is that the states of all systems at each level must be determined from simple logical operations on systems previously defined. Once the component systems have been defined the network must be defined. This is probably best accomplished in two steps because two kinds of information generally flow through the system network. In some cases the state of a given system is said to directly constrain the states which another system can occupy whereas in other situations the information that flows to a system only influences the decisions which specify the state. The two networks should be considered separately. The graph of the first type, the constraint network, should be constructed and the constraints themselves identified. After this has been accomplished the various systems which are decision loci are clearly identifiable and the decision influence network can be established.

In the next section these modelling ideas are applied to the problem of representing discrete control behavior in a simulated anti-aircraft artillery system.

3. DESCRIPTION OF THE AAA SYSTEM AND THE EXPERIMENT

The system which served as the focus of the study was a man-in-the-loop simulation of an anti-aircraft artillery (AAA) installation. This system consisted of a mock-up of the operators' consoles, including the major controls, switches, and displays; plus the computing equipment required to drive the displays, record data and generally simulate the AAA system and its environment. This particular simulator required a three person team consisting of a range tracker, an angle tracker and a commander. Basically the system consists of two optical sighting systems (left optics, right optics), a radar system with separate displays for the angle operator and range operator, a gun servo system which positions the guns as a function of tracking commands, and a lead angle computer. There are also a variety of switches and controls devices used to control and coordinate the activities of the system. The simulator is located at the Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base and the experiments which provided the data used in this analysis were performed by AMRL personnel.

The experiment was a highly stylized simulation of the AAA task. Subject teams were required to search for, acquire, and track simulated targets and to try to maximize the hit score attained on each target. The teams were free to select any system mode of operation at any point in time. Teams were asked to perform simulated "missions" which consisted of a sequence of 23 targets which they were required to acquire and track. Targets were presented one at a time and the target trajectory ran its full course before a new target was introduced. There was also a period of time between targets during which no target was present. The subject teams, however, remained actively involved in the search for the next target during this interval. Subject teams were given the approximate coordinates of the next target to maximize the

probability of detection. This information simulated the role of an early warning system. Two of the three subject teams completed a total of 23 missions and the third team completed 22.

Four distinct trajectories were used. One was considered a simple "non-threat" type of trajectory, the other three were considered "threat" trajectories by the experimenters. Six disturbance pattern types were imposed on the three easiest trajectories and five were imposed on the most difficult one. The net result was 23 distinct experimental conditions made up of disturbance type, trajectory combinations. Each experimental condition was presented once during each mission (hence 23 targets/mission). Presentation of the conditions was randomized from session to session. Also, data collection was grouped into blocks of four missions for each of the three teams. After each block was completed, the parameters which controlled the onset and duration of the disturbance conditions were modified to prevent learning of the disturbance patterns.

Data were collected in time-series format for every simulated mission in the experiment. The time set of the data collection is the mission time set which means that a complete running record of all measured variables was collected from the beginning to the end of a mission. The data collected included discrete status indicator type information, certain continuous tracking information, and header information to indicate teams and trajectories. All data were collected at a 30 HZ sampling rate. The discrete data collected consisted of all switch settings, and status display states, plus certain variables intended to provide information about the activities and performance of the team.

4. DECOMPOSITION OF THE AAA SYSTEM

The operators of the AAA system together must make decisions which determine which activities are to be engaged in and which mode of operation is to be used at each point in time. Their individual activities and decisions must be coordinated if the system performance level is to be maximized. Such coordination is achieved only if certain information flows through the system and each operator performs his tasks accurately and in a timely manner. The model of the AAA system used for analysis of the experimental data was carefully structured to capture the coordination and communication requirements as well as quantify individual task performance.

The complexity of the total system makes it necessary to decompose the system into a number of smaller, less complex systems which are responsible for certain specific tasks or serve as the information transfer points necessary for coordination. These smaller systems are placed in a network in which the systems themselves are the nodes and the arcs the communication links. The network serves as the model of

the team of operators working together on the discrete control problems imposed by the AAA system.

The network is probably best viewed as a related set of internal models. Certain parts of the network are best thought of as the internal models used by specific team members. But, other parts, particularly those which provide common communication points, are best thought of as models shared by two or more of the members. Through training and experience the team members learn the overall mission objectives and they learn what information must be shared, hence developing at some level a common representation of the problem and the system. Decisions made by the operators are presumed to be based on the state of these internal models. Furthermore, the decisions represent the desire to change the state of some component system thus enabling or disabling the occurrence of other events and decisions.

A brief overview of the specific systems used and some rationale for their construction is given below. More details are available in [1].

Several levels of analysis and types of decomposition are needed to construct a representation of a system as complex as this AAA system. The required decomposition takes place along three, not necessarily independent, dimensions:

1) Variable Type
   a) exogenous
   b) endogenous

2) Level of Analysis or Description
   a) primitive component
   b) major components
   c) functional systems
   d) coordination/communication systems
   e) management/command systems

3) System Type
   a) interface or information feedback
   b) decision controlled
   c) event controlled.

In terms of decomposition by variable type, the trajectory number, the target position system, and the disturbance system are the only exogenous variables. These systems are used to provide any target specific information used by other systems in the model and to explain any trajectory specific behavior. All other variables in the system are endogenous and characterize behaviors and decisions produced by the team members in response to the presented targets.

A hierarchical system involving the five levels of analysis listed above was used to establish the elements of the discrete control model. As one moves from the top of the list to the bottom, the view of the AAA system becomes more global and systemic. At the lowest level the perspective is that of individual controls and switches, at the top the perspective is that of overall mission objectives. Upper levels define abstract, less detailed views of the system; lower levels fill in the details. By moving from the top level to the bottom, any question about system performance and operation can be answered.

Level 1, the primitive component level consists of twenty simple systems which correspond to the basic switches and controls used during the experiments. The primitive components are listed in Table 1. These items are the decision or event controlled elements in the system. The state of each component listed in Table 1 must be established by some system higher in the system hierarchy.

Table 1

Primitive Components

Component Name	States
Antenna Horn	Search, Track
Automatic Circular Scan	Fast, Slow
Radar Mode	Circular Scan, Sector Scan, Manual, Automatic
Gun Servo Mode	Semi-Auto, Automatic
0^{o} Lead Enable	On, Off
Lead Enable	On, Off
Mode Switch	Mode I, Mode II
Computer Shunt	On, Off
Data Ready Indicator	On, Off
Coolant	On, Off
Trigger	On, Off
Display in Use	Optics, Not Optics
Azimuth Tracking Control	Rate, Position
Elevation Tracking Control	Rate, Position
Range Tracking Control	Fine, Coarse
Sight Selector	Left, Right
Sight Magnification	2x, 6x
Sight Filter	Clear, Neutral, Yellow
Lower Barrels	On, Off
Upper Barrels	On, Off

The system component level contains two types of systems; distinct system components such as the lead angle computer, and pseudo components defined by grouping certain primitive components which are manipulated together. The level two components are listed in Table 2. Complete state assignment rules are given in [1].

The sight system established the physical configuration of the optical sighting mechanism. The sight selector system is considered separately because of its importance in certain modes of operation (to be discussed when level four systems are discussed).

The range control control component is obvious. The gun configuration system defines precisely which barrels are enabled at any point in time and hence determines the maximum rate of fire possible. The angle track controls component establishes the exact configuration of azimuth and elevation controls.

The gun servo enabling network is a pseudo component. The state of this component establishes how tracking information flows through the system to the input of the gun drive mechanisms. Such flow can be disabled, fully enabled, or standby.

Table 2

List of Level 2 and Level 3 Systems

Level 2 System Components	Level 3 Functional Systems
Sight System	Fire Control Network
Sight Selector	Firing System
Range Control	Gun Directing System
Gun Configuration	Angle Track System
Angle Track Controls	Range Track System
Gun Servo Enabling Network	
Radar Antenna Drive	
Computer	

The radar antenna component defines the physical mode of operation of the mechanism which controls the motion of the antenna and the radar beam characteristics. State 1 is auto track which means that the track or narrow beam is in use and the range signals are under automatic control. The angle signals may also be under automatic control depending on the state of upper level systems. State 2 is manual tracking and states 3 through 6 are the various search modes which the team may use.

The computer is a physical component with three states of interest: standby, settling, operating. Settling refers to the period of time after which the computer is put in use but before a solution is reached. Operating refers to the period during which a lead angle solution is available.

The AAA system has other components, for example the A scope display used by the range operator, but none of these additional components have more than one mode of operation which is of operational significance. The components described above are precisely those which can potentially be used in multiple ways and which reflect the decision making activity of the team members.

Level 3 in the hierarchy is used to abstract out five major systems which perform the several functions which are prerequisite if the system is to meet engagement and mission objectives. These five are also listed in Table 2.

The states of the fire control network are defined to be locked, data enabled, fire enabled. When the fire control system is in the locked state the guns cannot be fired. When in the data enabled state, tracking data are available and with appropriate action by the angle operator the guns can be fired. When the fire enabled state is entered, the guns can be fired at any time. The firing system states are simply firing and not firing. The firing state is entered whenever the trigger is depressed, but firing actually occurs only if the fire control network is in the fire enabled state at the time. In other words, these systems must be coordinated for the overall system to function properly.

The gun directing system characterizes the status of the gun drive mechanism. The states are defined to be standby, 0^{o} lead tracking, and lead tracking. When the gun directing system is in the standby state the guns are not in motion. When in the 0^{o} lead state the input to the gun drive servomechanisms comes directly from the angle tracking system. In the lead tracking state the guns are driven by the output of the lead angle computer. The proper state for this system at any point in time depends on a number of factors which are established by the state of other higher and lower level systems.

The angle track and range track systems must provide the target state data which directly or indirectly drive the gun directing mechanism.

The angle track system states are defined to be optics auto, optics manual, radar auto, and radar manual. These define whether or not the angle operator is monitoring the PPI radar display or using one of the optical sighting systems and whether the angle track data available at the time is produced by automatic control or by the angle operator himself via manual control. The range tracking system, although logically a distinct system, has no autonomy in terms of discrete control. There is only one display and whether the mode is automatic or manual is completely controlled by other systems. The details of the range tracking system, therefore, need not be considered further for discrete control modelling.

Clearly, the systems which form the functional systems level of the hierarchy partition into three distinct groups: range and angle tracking; gun directing system; and fire control network and firing system. These groups define the three major functions of the AAA system: tracking targets, aiming the guns, and firing at the targets. Obviously, each of these systems must function properly for the mission objectives to be met.

Level four in the hierarchy is defined to be the communication/coordination level. The only system residing at this level is the engagement status system. This system is best thought of as a communication center through which information about the current activities of the system is passed. This information is then appropriately distributed and other system activities are enabled or disabled accordingly. The states of the engagement status system are: search, manual track, settling, and valid data. These states define the various conditions the system can be in from the beginning to the end of any single engagement. This system, together with other systems soon to be discussed, establishes whether or not things are progressing normally.

The fifth and highest level in the hierarchy is the management/command level. This level contains one system, the tactics system. This system is the locus of information and decisions concerning basic modes of operation. The tactics system, has five states: Normal Mode 1, Normal Mode 2, Mode 4, Emergency Mode 1, and Emergency Mode 2. Mode 1 refers to full automatic operation during tracking. That is, azimuth, elevation and range tracking data are all under full automatic control once the auto track mode (settling or valid data states of the engagement status system) is entered. The guns are directed by data from the lead angle computer in this mode. In Mode 2, only range data are placed under automatic control when the auto track mode is entered. Angle data are produced by manual tracking. The guns, however, are directed by the lead angle computer.

The emergency designation refers to fire control rather than tracking. In the emergency modes the computer shunt is turned on so that the guns can be fired whether or not the lead angle computer has reached a solution.

Mode 4 operation is a full manual mode in which the radar system is not used and the gun driver mechanism is slaved to the angle tracking output. This mode is

functional only if the angle track controls are in the rate mode (State 4). Furthermore, the only display which produces meaningful data in this case is the right optical system.

The state of the tactics system determines how the activities of the major functional systems will be carried out once the engagement status reaches the settling and valid data states. Further, it determines whether the guns can be fired prior to the valid data state. Finally, if mode 4 is selected the normal constraints imposed by the engagement states system are overridden but additional constraints must be imposed on component level systems if the system is to function properly.

The only additional systems which must be defined under decomposition by system type are the interface or information feedback systems. There are four such systems: Tracking Performance, System Performance, Ammunition Balance, and Mission Status.

The tracking performance system provides feedback about the quality of tracking. It is hierarchically defined in the sense that angle tracking errors are deemed more important than range track errors. The states of the system are: no target on any display; angles locked; angles OK, range locked and track OK. No target on display can occur if the tracking error is very large, or if there is no target to track. Angles locked is any case in which azimuth or elevation error is sufficiently large that the automatic tracking system cannot function. Range locked is a similar condition for range tracking error.

The system performance feedback systems attempts to capture some information about overall system performance. The states of this system are: no data, off target, on target. As designed this measure is a very local measure of system performance. It would be desirable to have a more global measure, but implementation problems prevented the use of such variables during this analysis. This system does, however, provide information about time on target, time off target and similar data. Clearly, if one or more major functional systems is not performing adequately, system performance state 3 will not be occupied.

The ammunition balance system determines the relative number of rounds in the upper and lower magazines. In the absence of other information, these data can be used to manage the use of ammunition resources.

The mission status system is used to assess overall ammunition resources with respect to the requirements of the remaining portion of the mission. The state of this system establishes whether or not special ammunition control (i.e., special concern with firing control) is needed if the mission is to be completed without depletion of resources.

These four systems provide the several systems in the discrete control hierarchy with information about local and global performance. This information, particularly any change in state, is used in part to determine if control actions are required.

5. THE COORDINATION AND DECISION CONDITIONING NETWORKS

Thirty-nine simple systems were defined during the process of decomposing the AAA system. These systems are the nodes in the networks which are the discrete control model. These structures organize the available knowledge about the AAA system and the discrete control tasks required for its operation. The objective in this section is to display these networks and examine some of their properties.

Although thirty-nine is a fairly large number of system elements to consider in a model of this type, each system is quite simple. No system has more than seven states and most have only two or three. Furthermore, it will be argued that the states occupied by these systems at any point in time are controlled by a fairly small number of decisions.

Two types of information must flow through the network. First, constraint information or direct control information and second, conditioning or influence information. The first type is said to actually cause specific state transitions to occur in other systems. Such transitions can be deterministic in that a specific state is occupied after the transition, or they can be non-deterministic in which case the new state is required only to be a member of a specific set. Conditioning information on the other hand does not directly constrain behaviors. Rather, it provides information to a given system about the state of other systems and this information may influence state transitions in the system receiving the information. Any state transitions which take place in this case are the result of a discrete control decision and this decision is based in part on the conditioning information in force at the time.

Several systems are controlled both by external systems and by internal decisions, depending on the situation. In specific situations this type of system's actions may be constrained or controlled by some other system in which case it is directly controlled. But, in other cases such constraints are relaxed and the behaviors of the system under question are decision controlled. This is one of the mechanisms by which overall coordination of the system is achieved and it is also a reason for structuring the system in a hierarchical fashion.

Figure 2 is a network diagram which shows the information sources and the receiving systems in the control/coordination network. The arcs (links) in this network should be thought of as communication channels through which the state of the originating node is made known to the receiving node. The state transitions which occur, or which are enabled to occur, in the receiving system are functions of the state of the originating system.

Space limitations prevent displaying the transition graphs of each system in Figure 2. The interested reader is referred to [1]. The general form of these diagrams is shown in Figure 3, which is the graph for the gun directing system. The elliptical

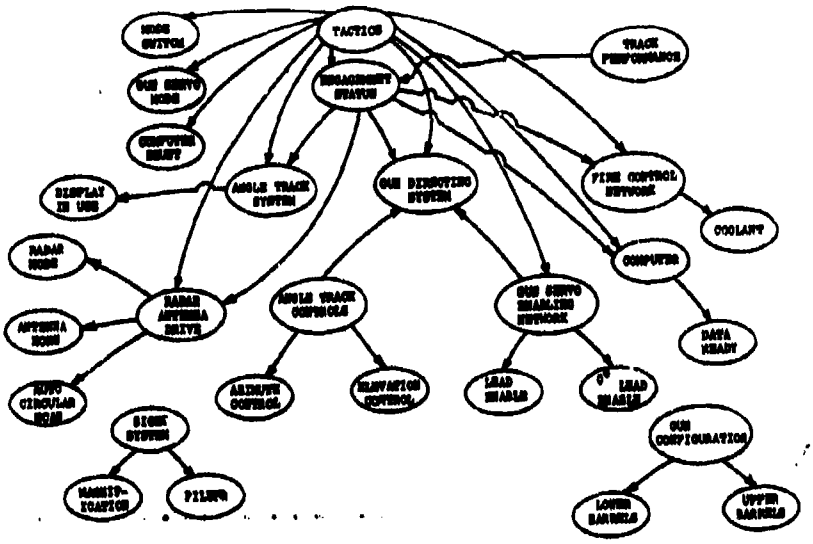


Figure 2.--Control/Coordination Network

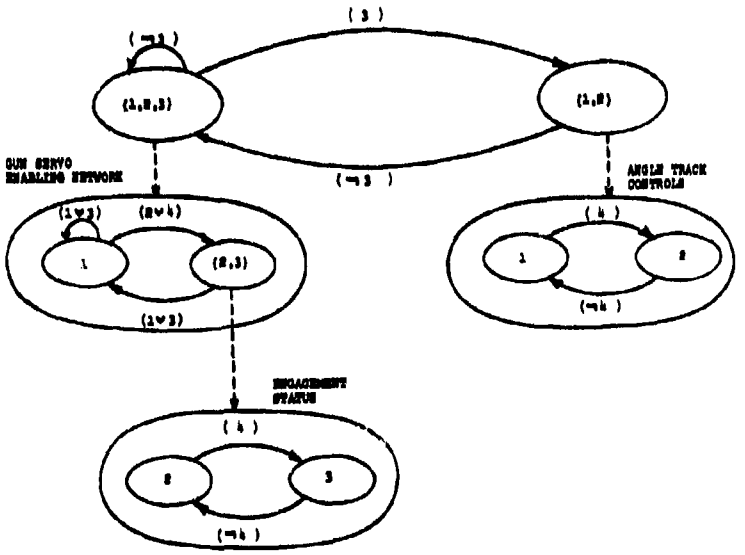


Figure 3.--Gun Directing System State Transition Diagram

figures represent states or sets of states. The description consists of the numerical value or set of values assigned to the node. State transitions are represented by solid lines connecting states and the input conditions which cause the transition to take place are defined by the bracketed symbols displayed on each transition arc. These symbols define logical expressions formed from the possible input values and a transition occurs when the appropriate logical expression is "true." To explain this more completely, any time the value of an input variable changes an event is said to occur and this event leads to a state transition in the system receiving the input. The transition which takes place depends on which logical expression is true at the time. Some nodes do not display transitions from themselves to themselves. These correspond to cases in which any input event will produce a transition out of the node in question.

The primary input variable, in this case tactics, explains transitions on the upper parts of the graph. The effect of secondary and other inputs is shown in those parts of the diagram connected to the primary via dashed lines. Secondary inputs are always used to provide a more detailed explanation of information in the primary diagram. They explain which state or states of the many allowed under the primary condition are actually occupied. A node connected to a primary node with a dashed line (a secondary conditioning variable) should be viewed as a more detailed representation of the primary node.

Several other important properties of the discret control model can be inferred from the coordination/control network and the corresponding transition graphs. First of all, any system for which all maximum resolution nodes contain single states (e.g., the gun directing system) is completely controlled by external sources. These systems for the most part are the lower level primitive components. The second class of systems is that for which one or more maximum resolution nodes is a set of states. Such systems are, at least under some circumstances, partly decision controlled. The third major class of systems, which with one exception do not appear on Figure 2, is event controlled systems. These are the information and feedback systems which interface the finite state systems with the various sources of continuous data. Lists of all three system classes are given in Table 3.

Of the systems controlled by decisions, seven are effectively controlled by external decisions in the sense that only in specific cases, usually indpendent on the tactics state, are they decision controlled. Of these, mode switch, automatic circular scan, antenna horn switch, 0^{o} lead enable, and lead enable, come under decision control only in cases in which their state is of no consequence. Generally, no state change would be made and these systems would remain in the state occupied prior to the occurrence of the event which placed them under decision control. The computer shunt is under decision control only when the tactics system is in state number three. In this situation the firing system cannot be operated unless the operator places the shunt in the on state. Hence, the computer shunt in actuality is constrained to be in state one if the system is to operate in these circumstances. The coolant system is under decision control only when the fire control network is in the locked state. The coolant

Table 3

Breakdown of Systems By Type of Control

Controlled By Internal Decisions	Controlled By External Decisions	Controlled By Exogenous Events
Tactics	Gun Directing System	Tracking Performance
Engagement Status	Computer	System Performance
Angle Track System	Gun Servo Mode	Ammunition Balance
Fire Control Network	Data Ready Indicator	Mission Status
Gun Servo Enabling Net	Display in Use	Target Position
Radar Antenna Drive	Radar Mode	Disturbance
Mode Switch*	Azimuth Control	
Computer Shunt*	Elevation Control	
Coolant*	Sight Filter	
Automatic Circular Scan*	Magnification	
Antenna Horn Switch*	Lower Barrel	
0^{o} Lead Enable*	Upper Barrel	
Lead Enable*		
Sight System		
Gun Configuration		
Sight Selected		
Range Control		
Firing System		
Angle Track Controls		

*Denotes systems which are effectively external decision controlled. See discussion in text.

state will not influence system performance until the fire control network is unlocked, and in that case coolant is controlled from the fire control network.

If the above systems are removed from the list, twelve systems remain in the decision controlled column. These can easily be partitioned in terms of importance.

The tactics system is obviously a key element. It interacts with ten other systems and it is the key element in establishing the system configuration. Engagement status is also a key element and it provides information to five systems. The angle track system, fire control network, gun servo enabling network and firing system follow in terms of impact on system configuration and overall performance. The remaining systems, although important, provide alternative means of accomplishing the same tasks and they probably have a lesser impact on total system performance. The decision flexibility in the radar antenna drive, for example, is in establishing the specific mode of automatic search. The major activity would be defined at a higher level.

Some general observations about system coordination can be made at this point. The coordination problem faced by the AAA team members might be defined roughly as follows: to direct each major subsystem into the proper state for each phase of an engagement. What is deemed the proper state will depend on the mission status, resources available and the characteristics of the target.

The network described above clearly illustrates a number of coordination activities. Specifically, the selection of the tactics state defines some major parameters which determine the configuration of the system and also the way in which the tracking phase of an engagement is to be carried out. For example, if the tactics system is placed in state three, the system is greatly simplified and the angle operator is responsible for manually finding and tracking any targets. The radar system, range tracking system, computer and most of the displays are of no interest. The communication of tactics information to the appropriate system elements then defines the set of states which those elements can use and thereby constrains behaviors to be consistent with the objective as defined by the tactics system.

The second point which should be made about coordination concerns the engagement status system. Whereas tactics determines the basic system structure and establishes what activities take place, enagement status provides the vehicle for coordinating the time phasing of these activities. In rough terms, the enagement states define what each system should be engaged in at a given time. Engagement status is the system through which the focus of control changes as the engagement evolves. During the search phase the focus in in the angle track system and associated subsystems. The status of all other systems is of very little concern during this time. During manual track the focus includes the angle and range tracking systems. During the settling phase, the focus of control is switching from the tracking systems to the gun directing systems and fire control networks. Once the valid data state is reached the focus is in

the fire control network and firing systems. During this phase the other system components are involved primarily in monitoring activities, trying to determine whether or not performance is satisfactory.

The heterarchical nature of the system is quite clear given the above perspective. Tactics sets some major constraints and unless a change is needed control flows to engagement status which in some sense directs control at the appropriate time to the tracking systems, gun directing, and fire control systems. A given system retains control until its task is complete or a lower or high level system intervenes and takes control for some reason. When a given system is the focus of control, the various subsystem which define it are active. The states of these subsystems are manipulated to accomplish the task. When a system is not the focus of control, its subsystems are much less active and generally exhibit no state change behaviors.

Errors and mistakes can also be described in terms of this network. The above discussion is based on the assumption that the operator or operators responsible for a specific activity were in fact prepared to carry it out. If control is given to a particular system and the operator whom this system represents in the specific situation fails to perform, he in essence has failed to accept control. This presumably would be detected and corrected at some point, but it certainly represents a deviation from the design condition and from standard procedure.

A second possible source of error exists in the class of systems which were called effectively decision controlled. Most of these systems have a nominal or preferred state and if for any reason the system state is changed during a period in which it is inactive, this might not be immediately detected when the system next becomes active. The operators would have to detect a problem and diagnose the source before making corrections and if the system causing the problem happens to be one whose state is seldom changed, this could take some time.

In summary then, the coordination/control network shown in Figure 2 defines the architecture of the discrete control system. It defines what information flows through the system, how activities and behaviors are enabled and disabled and they show how the focus of control is passed from one major system element to another. Furthermore, possible sources of error can be identified. These include the failure of an operator to accept control when it is passed to him and failure to detect an improper system state.

From Table 3 it was determined that the important decision controlled systems are the following: tactics, engagement status, angle track system, fire control network, firing system, gun servo enabling network, sight system, sight selected, range control, angle track controls, and gun configuration. Of these several can be removed from consideration because data analysis showed that the subject teams did not use the available alternatives. Tactics, for example, showed almost no decision activity and was set at emergency mode 1 by all teams. Gun configuration can be eliminated because

all decision activity was routine and not a major factor in resource management. Of the remaining systems, the overall behavior of the teams can be described in terms of four systems: engagement status, angle track system, fire control network and firing system.

After considerable testing, it was determined that the information needed to condition engagement status decisions is provided by track performance. Transition matrices and time summaries for teams one and two are given in Tables 4 and 5. As would be expected, transitions during the track OK situation are basically the same for both teams. Team two, however, generally shows a longer time in state than team one in this situation. They managed to stay in the valid data state for about 800 time frames which is roughly twice as long as team one.

Transition patterns during the no target phase are also different. Team two preferred to go into manual track from search if a change was made. Team one, on the other hand, actually started the lead angle computer a fair number of times in the no target situation. These transitions, however, took place after about 100 seconds without a target (3000 frames) and, therefore, they may correspond to cases in which the target was not detected.

The matrices for the angles locked case also show team one's reluctance to use manual track and a preference for transitioning from search directly to the settling state, state three. In other words, they preferred to try the computer even though tracking errors were large. The entries in row three show the same tendencies.

Angle track system activity as a function of tracking performance is given in Tables 6 and 7. Both teams in the no target situation show a preference for the radar-auto state, state one. Apparently search was accomplished with the radar system in an automatic mode (sector search). Most transitions from state one were to state two, the radar-manual state. This transition signals the start of manual search. Transitions from state two in the no target situation were most often to the optics-manual state for team one which means they were almost always making a display change, probably in an attempt to find the target. It is interesting to note that both teams almost always (probability .979) transitioned from state two to state 4, optics manual, in both the angles locked and the range locked cases. They both also showed a high probability of exiting the radar-auto state. These imply a very strong preference to complete the target acquisition phase using the optical sighting system. Furthermore, almost all transitions from state four were to state two in the cases where angle error and/or range error were locked. This means that the activity during manual acquisition of the target consisted of display changes. The track OK matrices show the very definite tendencies to get into auto track and use the optical display system.

The conditioning variables for the fire control network are the angle track system and the engagement status. Nearly all team two activity took place in engagement status states one or four (which are search and valid data) and angle track system

Table 4A

Team 1
Engagement Status Transition Matrices

Track Performance	Transition Matrix				Transition Count
1 (No Target)	0.000	0.407	0.584	0.009	113
	0.924	0.000	0.076	0.000	131
	0.615	0.038	0.000	0.346	78
	0.907	0.093	0.000	0.000	97
2 (Angle Error Locked)	0.000	0.295	0.705	0.000	88
	0.000	0.000	0.000	0.000	0
	0.900	0.000	0.000	0.100	40
	0.800	0.700	0.000	0.000	5
3 (Angles OK, Range Locked)	0.000	0.530	0.470	0.000	215
	0.722	0.000	0.278	0.000	72
	0.000	0.000	0.000	0.000	0
	0.000	0.000	0.000	0.000	0
4 (Track OK)	0.000	0.000	1.000	0.000	128
	0.000	0.000	1.000	0.000	25
	0.081	0.003	0.000	0.916	298
	0.873	0.127	0.000	0.000	197

Table 4B

Team 1
Engagement Status Time In State Data

Track Performance	Transition	Avg. Condition Time (Frames)	Standard Dev.	Count
1	1-2	816	1181	46
1	1-3	3090	3091	66
1	2-1	1016	1166	121
1	3-1	103	184	48
1	4-1	535	298	88
2	1-2	2060	2020	26
2	1-3	3712	4491	62
2	3-1	50	19	36
3	1-2	2212	2534	114
3	1-3	4402	3557	101
3	2-1	1088	809	52
3	2-3	1950	940	20
4	1-3	3872	3726	128
4	2-3	2580	904	25
4	3-1	94	133	24
4	3-4	101	0	273
4	4-1	378	222	172
4	4-2	546	209	25

Table 5A

Team 2
Engagement Status Transition Matrices

Track Performance	Transition Matrix				Transition Count
1	0.000	0.892	0.108	0.000	93
	0.958	0.000	0.042	0.000	261
	0.455	0.364	0.000	0.182	11
	0.667	0.167	0.166	0.000	6
2	0.000	0.955	0.027	0.018	111
	0.000	0.000	0.000	0.000	0
	0.250	0.750	0.000	0.000	4
	1.000	0.000	0.000	0.000	3
3	0.000	0.805	0.177	0.018	169
	0.738	0.000	0.250	0.012	84
	0.000	0.000	0.000	0.000	0
	1.000	0.000	0.000	0.000	3
4	0.000	0.000	1.000	0.000	152
	0.045	0.000	0.955	0.000	22
	0.033	0.000	0.000	0.967	241
	0.830	0.157	0.013	0.000	223

Table 5B

Team 2
Engagement Status Time In State Data

Track Performance	Transitions	Avg. Condition Time (Frames)	Standard Dev.	Count
1	1-2	804	1132	83
1	1-3	3978	2151	10
1	2-1	518	625	250
2	1-2	2567	3245	106
3	1-2	2680	3035	136
3	1-3	5076	4854	30
3	2-1	425	373	62
3	2-3	835	457	21
4	1-3	5121	4604	152
4	2-3	779	644	21
4	3-4	101	0	233
4	4-1	809	296	185
4	4-1	807	267	35

Table 6

Team 1
Angle Track System Transition Matrices

Tracking Performance	Transition Matrix				Transition Count
1	[0.0000	0.716	0.2840	0.0000]	455
	[0.1361	0.000	0.0005	0.8634]	2035
	[0.6310	0.000	0.0000	0.3690]	274
	[0.0000	0.945	0.0550	0.0000]	1574
2	[0.0000	0.812	0.1730	0.0000]	16
	[0.0210	0.000	0.0000	0.9790]	373
	[0.0530	0.000	0.0000	0.9470]	38
	[0.0000	0.839	0.1610	0.0000]	336
3	[0.0000	1.000	0.0000	0.0000]	2
	[0.0160	0.000	0.0000	0.9840]	1134
	[0.0000	0.000	0.0000	0.0000]	0
	[0.0000	0.927	0.0730	0.0000]	1427
4	[0.0000	0.261	0.7370	0.0020]	414
	[0.8460	0.000	0.0000	0.1540]	26
	[0.7730	0.014	0.0000	0.2130]	497
	[0.0000	0.064	0.9360	0.0000]	140

Table 7

Team 2
Angle Track System Transition Matrices

Tracking Performance	Transition Matrix				Transition Count
1	0.000	0.627	0.372	0.001	825
	0.594	0.000	0.000	0.406	1093
	0.744	0.000	0.000	0.256	391
	0.000	0.957	0.043	0.000	507
2	0.000	0.912	0.088	0.000	34
	0.021	0.000	0.000	0.979	188
	0.333	0.000	0.000	0.667	9
	0.000	0.973	0.027	0.000	111
3	0.000	1.000	0.000	0.000	17
	0.033	0.000	0.000	0.967	546
	0.000	0.000	0.000	0.000	0
	0.000	0.936	0.064	0.000	516
4	0.000	0.437	0.560	0.003	359
	0.833	0.000	0.000	0.167	18
	0.784	0.114	0.000	0.102	324
	0.000	0.031	0.969	0.000	163

states three or four (optics-auto and optics-manual). That is, team two almost always used the optics when firing. Team one showed an unexpected amount of activity with the angle track system in the radar-auto state. This suggests that they may have on occasion prepared to fire before switching to the optical sighting system.

The condition variables for the firing system are the fire control network and engagement status. One interesting item was observed from the firing system analysis. Team one had a tendency when manually tracking (engagement status one) to try to fire the guns without first putting the fire control network into the fire enabled state. It appears that when involved with the tracking activities they sometimes forgot how the system worked and deviated from the standard procedures. Team two did not have this problem.

The decision conditioning network is the scheme by which the conditioning which was used in the above analysis is best described. This network which was empirically derived is shown in Figure 4. Essentially, tracking performance influences decisions made in the angle track and engagement status systems. Engagement status then influences the activities of the firing system and the fire control network as well as some of the lesser systems. Those systems shown in Figure 4 which are not connected to any other system are those systems which were either routinized or showed little activity. This simple network localizes essentially all of the decision making activity which was shown in the data. The conjunction of this network and the coordination/control network shown in Figure 2, together with the appropriate state transition graphs, are the discrete control model obtained from an analysis of the system and a comprehensive analysis of the data.

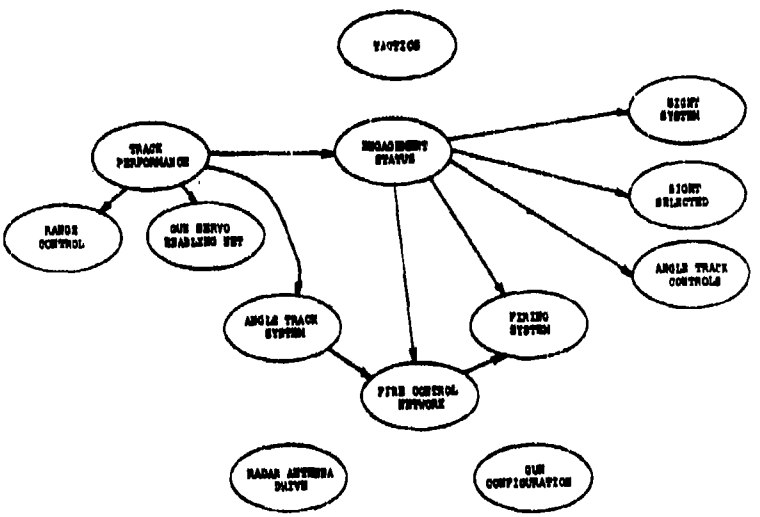


Figure 4.--Decision Conditioning Network.

6. CONCLUSIONS

Based on the analysis it is fairly easy to construct a scenario of the way in which the subject teams performed their tasks. First of all, search for targets was generally accomplished using the automatic sector search mode and the primary display was the PPI radar screen. In cases where the target was slow to appear some teams might occasionally switch to the left optics sighting system and less often to the right optical system. As soon as the target appeared the angle operator would switch to the left optics system and switch from the radar driven search mode to a manual mode for acquisition. During this period the range operator was switching his control back and forth between the coarse and fine setting. The commander was resetting the gun servo enabling system to state three and thereby moving the gun directing system from standby to 0^{o} lead tracking. Also during the acquisition phase and the early phases of tracking the angle operator determined which of the four trajectories the current target was following. This determined in part the strategy that he then followed.

Once the target was acquired the system was either put into auto track or, if the trajectory was particularly easy, tracking often continued in a manual mode. Team two used manual tracking for the most difficult highly maneuvering target as well. If manual tracking was used, firing started within about 20 seconds of acquisition (i.e., when the target came within range). In cases where auto track was used, firing started shortly after the fire control solution was achieved. Easy targets were disengaged after a few hits. The gun directing system was then generally disabled for a few seconds after the target disappeared from the screen. The angle track system was then reconfigured for search and the cycle started again.

Several interesting observations can be made about the performance of the teams. Team one had some problems with the fire control network interlocks when they used manual tracking. They forgot to enable firing before trying to fire. The other teams had no problem and team one had no problem in the automatic tracking cases. Team one may have become so involved in performing the tracking task that they forgot how the system worked. There were other pieces of evidence which showed that the teams infrequently made incorrect switch settings or failed to reconfigure the system quickly enough.

It was a surprise that no team used the six power optical system very much for tracking. Apparently the feedback provided by tracers reduced the need for precise visual information. It definitely seems that a style of tracking was used in this experiment which differed qualitatively from that used in simple tracking studies.

In general terms the commander's tasks were very trivial. Teams obviously learned the limited number of trajectories which were used and they keyed their actions to the trajectory. The attempt to introduce uncertainty via the disturbances or simulated countermeasures did not seem to have much impact. They may have delayed the start

of auto track for example, but they did not alter the basic patterns of behavior as represented by the various transition matrices of the discrete control model. The fact that every mission contained exactly the same number of trajectories greatly simplified ammunition management. The subject teams knew that they could and should go after all targets. They did not have to be selective or evaluate the threat potential of any target. There was certainly no risk associated with missing one and there was no significant scoring penalty.

The model which seems to best capture the various teams' performance is really a set of finite state systems organized into two networks, the coordination/control network (Figure 2) and the decision conditioning network (Figure 4). The systems included in these networks were established through a detailed analysis of the AAA system, its functions, and the tasks of the operators. By decomposing along several structural dimensions, and particularly by analyzing at several levels of abstraction, an effective and useful representation of the discrete control system was obtained.

In general terms this representation is a model of an organizational structure which the operators might use to reduce the apparent complexity of their task and generally achieve coordinated actions and acceptable performance. It is really just a structured representation of the available knowledge of the system and its functions.

The coordination/control network is basically hierarchical and reflects the constraints on lower level decision making activity imposed by upper level decisions. In the terms of the finite state systems representation, state transitions in lower level systems are disabled, enabled or constrained as a function of the state of upper level systems. The decision conditioning network establishes the information flow patterns which are needed to explain, at least in part, the decisions which are made (i.e., the state transitions which take place). The systems in the decision network are represented by generalized stochastic automata in which state transitions are conditioned by the information flowing into the system from other nodes in the network. The two networks in conjunction form a heterarchical system description in which decision making activity flows from one functional area to another as a function of the established constraints and the environmental situation.

In general terms then, the discrete control methods so far developed seem to have potential. They can be used to make sense out of complex systems and identify the key decision points. They can describe quite complex behaviors in terms of a relatively small number of decisions. The structure of the model is quite easy to understand and the individual finite state systems are all simple and intuitive. Grasping the overall view; i.e., all levels simultaneously, is more difficult and amount of statistical information which can be produced is overwhelming. These problems are minimized, however, if one restricts attention to only the one or two levels which are most important for a given question.

REFERENCES

1. R. A. Miller, Ohio State University Research Foundation, Ohio State University, "Identification of Finite State Models of a Human Operator," Final Report, AFOSR Grant No. 77-3152, March 1979.

ACKNOWLEDGEMENTS

This research was supported in part by the Air Force Office of Scientific Research, Grant #77-3152 and this support is gratefully acknowledged.

PROGRESS REPORT ON THE TELECONFERENCE CUM ELECTRONIC JOURNAL ON MENTAL WORKLOAD

Thomas B. Sheridan (1-110)
Massachusetts Institute of Technology
77 Massachusetts Ave.
Cambridge, MA 02139

ABSTRACT

This paper summarizes the first few months of activity in the National Science Foundation supported computer teleconference on mental workload, and the plans to produce on a trial basis an "electronic journal" papers for which will be written, edited, refereed, revised, and commented upon by the interested community entirely over computer terminals. Issues which seem to be of primary concern in the "mental workload field" will be mentioned, along with problems of running a computer-based teleconference cum journal on such as outlandish topic.

2

SESSION 2: CONSIDERATIONS FOR MAN-IN-THE-LOOP SIMULATIONS

Moderator: Dr. Sheldon Baron, Bolt Beranek and Newman

THE EFFECTS OF SIMULATION FIDELITY ON AIR-TO-AIR TRACKING

Billy R. Ashworth,
Burnell T. McKissick,
Russell V. Parrish

NASA Langley Research Center
Hampton, Virginia

ABSTRACT

NASA Langley recently sponsored development of an analytic model of the LaRC real-time digital simulation facility which incorporates the optimal control model for the human pilot. In order to use the model to evaluate simulator design tradeoffs, an in-simulator experiment was conducted to provide validation data for the "system model." The results of the in-simulator experiment and the system model predictions show a significant effect on pilot performance for the five simulation factors studied (i.e., integration algorithm, control loader bandwidth, g-seat cueing, computer iteration rate, and visual time delay).

THE EFFECTS OF SIMULATION FIDELITY ON AIR-TO-AIR TRACKING

Numerous tradeoffs between simulation fidelity and complexity (cost) must be made in the development of engineering requirements for man-in-the-loop simulation. NASA-Langley recently sponsored development of an analytic model (ref. 1) of the LaRC real-time digital simulation facility which incorporates a model for the human pilot (namely, the optimal control model). This total closed-loop model allows simulator design tradeoffs to be quantitatively evaluated.

In order to provide validation data for this model, an experiment was conducted in LaRC's Differential Maneuvering Simulator (DMS) to examine several "simulation fidelity" factors. Initially five factors were chosen to be examined in a pilot-in-the-loop tracking task. The factors chosen were: Computer iteration rate (16 and 32 iterations per second); integration algorithm (Euler and Adams-Bashforth); visual time delay (68 milliseconds and 12 milliseconds); control loader mechanization (digital or analog loop closures); and g-seat cueing (off or on). The iteration rates and integration algorithms are representative of those used in numerous simulation facilities. The values chosen for visual time delay, control loader mechanization, and g-seat cueing represent those available in the DMS for the model validation.

FACILITY DESCRIPTION

The DMS (fig. 1) is a wide-angle visual simulator which allows one-on-one aircraft interactions. The servos involved in projecting the visual scene for the pilots all have a matched second-order response with a 56 millisecond steady-state time delay. The cockpits are also equipped with wide FOV HUD's (25° x 35°) which present a computer-generated image to the pilot. The fixed-base cockpits are equipped with programmable, hydraulically actuated control loader systems, a g-seat, a helmet loader, and a g-suit drive to provide some motion induced forces for the pilot. For more complete information, see Reference 2.

In order to obtain the low visual delay, the target image and horizon line were presented to the pilot on the HUD (fig. 2). The visual delay of the HUD symbology is 12 milliseconds; therefore a 25 rad/sec, second-order filter with a 56 millisecond steady-state time delay was added to the display to obtain a response representative of the normal projection servos.

The cockpit control loader has been modified to allow analog loop closures as well as the more conventional digital force feedback. For information on the controller bandwidth reduction due to digital-loop closures, see Reference 3. The g-seat presents the normal acceleration cues to the pilot through the seat pressures (refs. 4 and 5). The aircraft dynamics programmed for this particular experiment represent a linearized version of the F-8 aircraft with capability for single or dual axis tracking with fixed range (ref. 1).

STATISTICAL ANALYSES AND EXPERIMENTAL DESIGN

Experimental Task

A single axis compensatory pitch tracking task was used for the experiment. A block diagram of the task is presented in Figure 3. A sum of 13 sinusoids with a fixed set of amplitudes and frequencies but randomly chosen phases (between -180° and +180°) was used to drive the vertical degree of freedom of the target aircraft. The phases were randomly chosen so that the test subjects would not learn the movements of the target. Table I presents the amplitudes and frequencies in the sum of sine waves. For the experiment, the target and pursuit aircraft were limited to the longitudinal degree of freedom only. Therefore, neither aircraft could roll, yaw, or change throttle controls (range). Both aircraft were trimmed about level flight at an altitude of 6,096 meters (20,000 ft) with a constant speed of 325 knots and a range of 182.88 meters. The pursuit aircraft used a standard backup reticle for the tracking reference. In order to "track" the target, the pursuit aircraft was required to keep the tail of the target in the center of the reticle.

Experimental Design

The factors of the design are denoted as follows: (A) integration schemes (Euler or Adams-Bashforth), (B) control loading (digital or analog), (C) g-seat (off or on), (D) integration step size (16 per second or 32 per second), and (E) visual delay (68 milliseconds or 12 milliseconds). The total number of simulation conditions for the 2^5 factorial design would be 32. When the number of test subjects (8) and the number of replicates (at least 5) are taken into consideration one ends up with 1280 as a minimum number of simulation runs. Since this number of runs would be too expensive, the 2^5 factorial design was fractionated and blocked. Figure 4 is a tree diagram of the experimental design chosen. First, the highest order interaction, ABCDE, was used to fractionate the 2^5 factorial in half. One-half of the 2^5 factorial or 2^4 conditions were not used. Next, the DE interaction was used to block the half that was used into two parts. Then, the CD and BE interactions were used to block the two parts into four parts each containing 2^2 conditions. Finally, two test subjects (pilots) were assigned to each block. This design reduced the number of simulation runs from 1280 to 160.

It should be noted that the interaction terms selected for aliasing through the fractionalization and blocking process were chosen based on predictions of non-significance from either system model results or prior in-simulator results.

Test Subjects and Procedure

Eight active F-106 pilots, stationed at Langley Air Force Base, were used as test subjects. None of these pilots had previous experience in a research simulator like the DMS. The pilots were randomly assigned to a block. The conditions inside each block were randomly chosen to determine the order in which the conditions would be performed.

A data collection period for each simulation condition took about 3 minutes per run. The first 20 seconds were used to phase in the sum of sines disturbance. The next 40 seconds were used for pilot stabilization and the final 2 minutes were used for data collection. In order to encourage the pilots to do their best, their RMS vertical tracking error was given to them after each simulation run.

Simulation sessions are usually allotted in 3-hour slots at LaRC's simulation facility. Previous experience suggested that 15 runs of the compensatory tracking task was about the maximum number of runs that could be made before the pilot's performance started to deteriorate because of fatigue. The majority of the 15 runs were taken up by the pilots' learning of the task.

Therefore, it was felt that two pilots could perform the experiment per simulation session, or one pilot could perform two different conditions (with proper rest between the first and second performance) during one simulation session. Normally, a pilot would "fly" the simulator with a randomly chosen condition until the RMS of his vertical error "asymptoted." Generally, this took from 5 to 10 replicates. Once his vertical error has asymptoted, the pilot would do five more replicates that were used as data.

Each pilot was tested at four different conditions. Since two pilots were assigned to each block (fig. 4), two pilots would be running the same four conditions. To further eliminate any learning/fatigue effects, the pilots within a block flew the conditions in reverse order.

Performance Measures

The RMS of the vertical height error, pitch Euler angle, normal acceleration, and pilot stick deflection were computed as measures of tracking performance. Also, the pilot describing function (pdf) was computed as a measure of performance. The method used to compute the pilot describing function is documented in Reference 6.

Statistical Analyses

Three types of analyses were made--analyses of variance, Student's t-tests, and F-tests (homogenuity of variance tests). The analyses of variance were performed on the RMS data and the amplitude ratios, phases, and remnants of the pdf's at each of the 13 frequencies of the sum of sinusoids. The t-tests and F-tests were performed respectively on the means and variances of the RMS data.

RESULTS

The RMS of the vertical tracking error was chosen as the primary performance measure. The means and standard deviations of the RMS tracking error for each level of each of the five factors (averaged over the other four factors, $n = 80$) are shown in Figure 5. The probabilities beside the higher of the two means for each factor represent the significance levels of a t-test on these means. The probabilities listed beside the $+\sigma$ points represent the significance levels of an F-test on the variances of each factor. The probabilities of 1.00 occur because of round off errors (0.99999 or greater) in the computer subroutine used to perform the statistical tests. The N.S.'s in the figure mean "not significant." It is easily seen from Figure 5 that the RMS tracking error is reduced by using AB-2 integration or 32 iterations per second or no visual delay (both mean and variance reductions are significant), or g-seat on or analog control loading (mean reductions, though not significant).

Table II presents the results of an analysis of variance for the five main factors on the raw RMS data of tracking error, as well as RMS of pitch (θ), RMS of normal acceleration (N_z), and RMS of pilot stick deflection for pitch control. The results show that the reduction in mean tracking error is statistically significant at the 99 percent level for integration algorithm, iteration rates, g-seat cueing, and visual delay. Control loader is significant at the 90 percent level. Other RMS errors are also significantly affected by the various fidelity factors.

The pilot's linear transfer function (pilot describing function) for each factor was also used as a measure of performance. The describing functions for the two levels of iteration rate are given in Figure 6 (averaged across each level of the other four factors, n = 80). The pilot describing function (pdf) was computed at the 13 sum of sine input frequencies and an analysis of variance was performed on the raw pdf data at each frequency. These results are shown in Tables III, IV, and V for the gain, phase, and remnant, respectively. All factors show a significant difference in gain for the lower frequencies. In addition, the three stronger factors (integration scheme, iteration rate, and visual delay) show significant differences at higher frequencies. There is also scattered but significant difference in phase for the stronger factors in the low and middle range of frequencies. The remnant is significantly different for integration scheme in the low and middle frequencies, for control loading and visual delay in the higher frequencies, and across all frequencies for iteration rate.

The data in the tables can be viewed in terms of the plot of pdf as shown in Figure 6. The differences in the pilot gains for the two curves (16 and 32 iterations per second) show the low frequency differences that are listed as statistically significant in Table III. The fact that the pilot gain is higher for the "better" condition (32 iterations per second) is true for all cases; therefore, any significant difference listed in Table III implies significantly higher pilot gain for the better of the two levels of the effect. The data trends for phase and remnant are inconsistent across the better/worse conditions, with reversals occurring from frequency region to region.

MODEL COMPARISON

The untuned pilot model was initially used to look at the effects due to visual delay (E), integration algorithm (A), iteration rate (D), and control loader loop closure (B). These initial results showed significant differences in RMS tracking error for each factor, with A, D, and E being the strongest, respectively. A worst (Euler integration, 16 iterations per second, and 56 milliseconds time delay) versus best (AB, 32, 12 milliseconds) case was defined for the three strongest effects (A, D, E).

These pre-experiment results are shown in Figure 7 along with similar results from the in-simulator experiment. Although the case differences of the "untuned" model were not as strong as the eventual simulator results, the trends are correct and do show a significant difference in the RMS tracking for the two cases.

The data in Table VI were generated from "tuned" model runs for each of the 2^4 combinations of fidelity factors (Factor B was averaged out for the best/worst comparisons). For example, the Euler mean was obtained by averaging over the eight runs that used Euler integration.

The earlier trend of the model being somewhat optimistic in the "worst" conditions is continued with the "tuned" model as shown in the first column, Table VI. However, when the observation noise/signal ratio was changed from -17 db to -11 db (for the Euler integration case at 16 Hz sample rate only) as shown in the second column of the table, the match with the experimental data is much closer. A plot of this final match is shown in Figure 8.

CONCLUSIONS

The five simulation factors identified and studied each had a significant effect on pilot performance in a simulated air-to-air tracking task. The pilot performance measures showed significantly better performance for Adams-Bashforth second-order integration, a 32 iteration per second sample rate, and a 12 millisecond total visual delay. G-seat cueing and analog control loader feedback also showed significantly lower errors from the analysis of variance results. The pilot describing function data revealed a higher pilot gain for all five of the factors at these "better" conditions.

These general results, with the exception of g-seat cueing, were predicted by the system model, as reported last year in Reference 1. The experimental data were more sensitive to the worse level of each factor than the original "untuned" model had shown. However, running the system model under the same conditions as the experiment and changing the observation noise for Euler integration at 16 Hz gave excellent comparisons to the experimental data.

The results of this study demonstrate the use of the simulation facility systems model to predict simulation fidelity effects. Tuning the model about local conditions from in-simulator results then allows the analysis of both hardware and software design increments about those local conditions. This tool for the analysis of flight simulation engineering requirements (as well as simulator constraints on pilot/vehicle performance) should prove extremely useful in further applications.

REFERENCES

1. Baron, S.; Muralidharan, R.; and Kleinman, D.: Closed Loop Models for Analyzing Engineering Requirements for Simulators. Bolt, Beranek, and Newman, Inc., Report No. 3718, May 1978.

2. Ashworth, B. R.; and Kahlbaum, W. M., Jr.: Description and Performance of the Differential Maneuvering Simulator. NASA TN D-7304, June 1973.

3. Parrish, R. V.; and Ashworth, B. R.: The Effect of Digital Computing on the Performance of a Closed Loop Control Loading System. NASA TN D-8371, December 1976.

4. Ashworth, B. R.: A Seat Cushion to Provide Realistic Acceleration Cues for Aircraft Simulator. NASA TMX 73954, September 1976.

5. Ashworth, B. R.; McKissick, B. T.; and Martin, D. J., Jr.: Objective and Subjective Evaluation of the Effects of a G-Seat on Pilot/Simulator Performance During a Tracking Task. Proceedings of the 10th NTEC/Industry Conference (Orlando, Florida), November 1977.

6. Shirachi, D. K.; and Shirley, R. S.: The Effect of a Visual/Motion Display Mismatch in a Single Axis Compensatory Tracking Task. NASA CR-2921, October 1972.

TABLE I.- PARAMETERS OF SUM OF SINES

Frequencies (rad/sec)	Relative Amplitudes (Without Scale Factor of 6.1 m)
0.245	1.150
0.540	0.747
0.933	0.319
1.424	0.121
2.013	0.511
2.896	0.022
4.074	0.009
5.547	0.004
8.001	0.002
10.946	0.001
16.248	0.0003
22.040	0.0001
32.094	0.00006

TABLE II.- ANOVA RESULTS FOR RMS PERFORMANCE METRICS

Effect	Metric			
	RMS Error	RMS θ	RMS N_Z	RMS Stick
Integration Algorithm (A)	***	***	***	
Control Loader (B)	*			**
G-Seat (C)	***	*		
Iteration Rate (D)	***	***	***	
Visual Delay (E)	***	***	**	

*** ⟶ ANOVA PROBABILITY >0.99

** ⟶ ANOVA PROBABILITY >0.95

* ⟶ ANOVA PROBABILITY >0.90

TABLE III.- PILOT DESCRIBING FUNCTION AMPLITUDE RATIO

Effect	Frequency												
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13
Integration Algorithm (A)	***	***	***	***	***	***			*	**	*		**
Control Loader (B)		***	***	***	*						*	*	
G-Seat (C)		**	***	*	**								
Iteration Rate (D)	***	***	***	***	***	*				**	***	***	***
Visual Delay (E)	*	***	***	***	***			***		***	*	***	*
		Low				Middle				High			

***→ANOVA Probability > 0.99
**→ANOVA Probability > 0.95
*→ANOVA Probability > 0.90

TABLE IV.- PILOT DESCRIBING FUNCTION PHASE

Effect	Frequency												
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13
Integration Algorithm (A)			***		***	***	*						
Control Loader (B)			***										
G-Seat (C)			*										
Iteration Rate (D)		***	***		*		**	*	*				
Visual Delay (E)	*			**	***			*					
					Low			Middle				High	

***→ANOVA Probability > 0.99
**→ANOVA Probability > 0.95
*→ANOVA Probability > 0.90

TABLE V.- PILOT DESCRIBING FUNCTION REMNANT

Effect	Frequency												
	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10	#11	#12	#13
Integration Algorithm (A)	***	***	***	**	*	*	**	**	*				
Control Loader (B)		*	*				*	**		***	***	***	***
G-Seat (C)													
Iteration Rate (D)	***	***	***	***	***	***	***	***	***	***	***	***	***
Visual Delay (E)	*			***				**	**	***	***	***	**
	Low				Middle					High			

***→ANOVA Probability > 0.99
**→ANOVA Probability > 0.95
*→ANOVA Probability > 0.90

TABLE VI.- AVERAGE MODEL PERFORMANCE AND STANDARD DEVIATION

		RMS VERTICAL TRACKING ERROR (FT)		
		Observation Noise/ Signal Ratio (-11 db)	Observation Noise/ Signal Ratio (-11 db for Euler at 16 Hz only)	Experiment
A	Euler	10.76 ± 0.74	12.69 ± 2.50	12.67 ± 3.19
	AB2	9.67 ± 0.68	9.67 ± 0.68	9.69 ± 1.59
D	16 Hz	10.72 ± 0.81	12.66 ± 2.60	12.57 ± 3.11
	32 Hz	9.69 ± 0.65	9.69 ± 0.65	9.79 ± 1.89
E	56 ms Visual Delay	10.59 ± 0.79	11.55 ± 2.33	12.26 ± 2.90
	0 ms Visual Delay	9.83 ± 0.83	10.81 ± 2.41	10.10 ± 2.52
B	Digital CL	10.49 ± 0.91	11.45 ± 2.43	11.35 ± 2.83
	Analog CL	9.93 ± 0.6	10.91 ± 2.36	11.01 ± 3.02
	Worst	11.69 ± 0.52	15.50 ± 0.46	16.27 ± 2.08
	Best	8.80 ± 0.28	8.80 ± 0.28	8.39 ± 0.97

NASA
L-71-8700

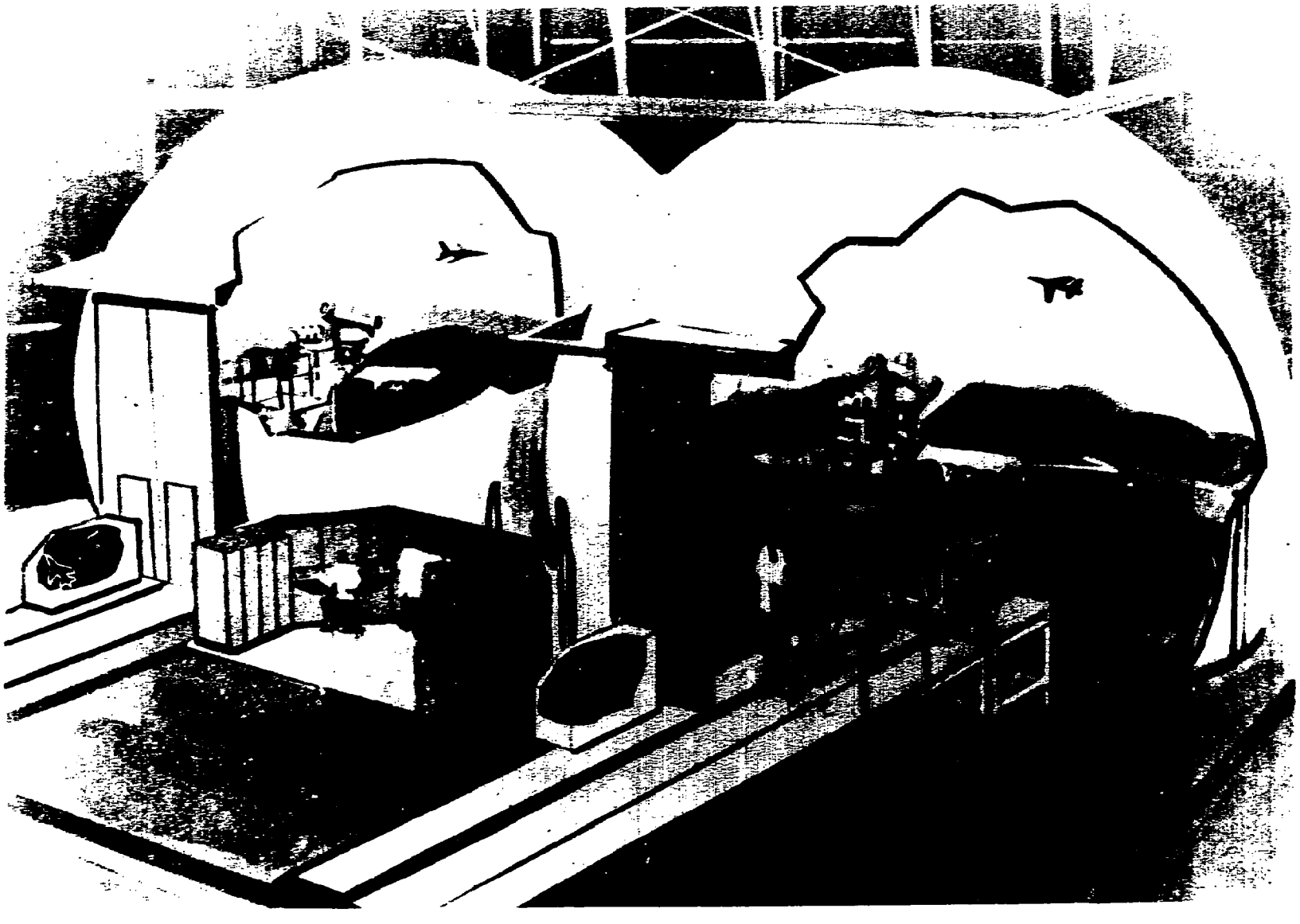


Figure 1.- Differential Maneuvering Simulator.

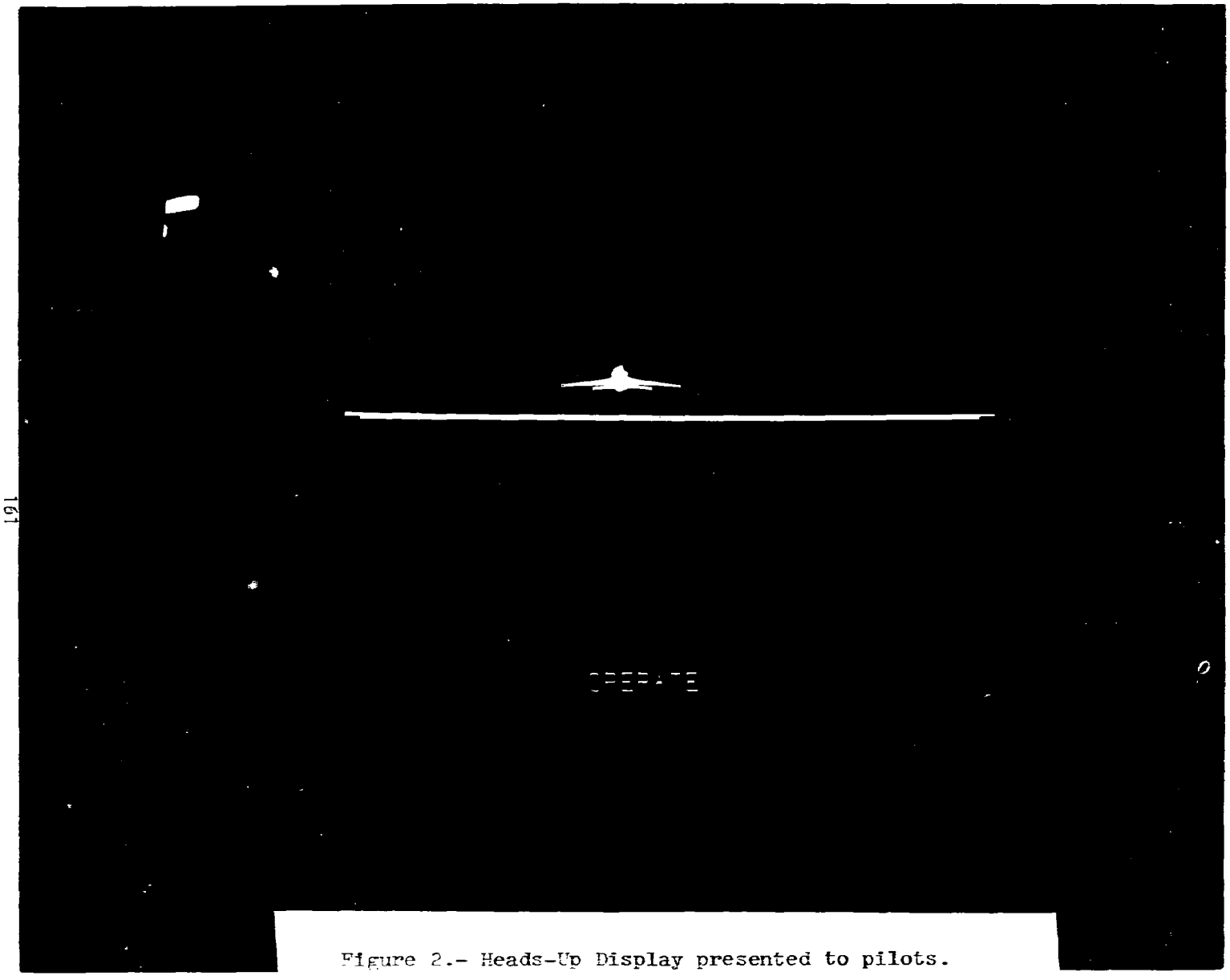


Figure 2.- Heads-Up Display presented to pilots.

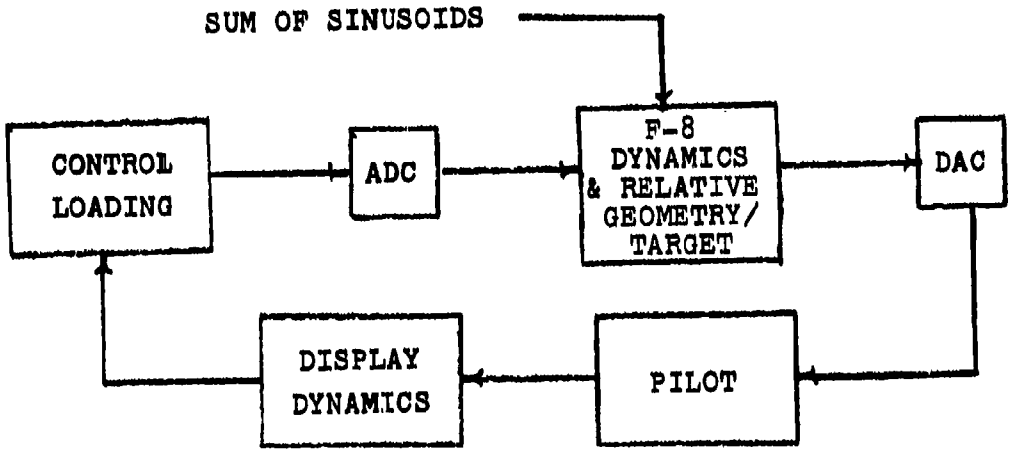


Figure 3.- Single axis compensatory tracking task.

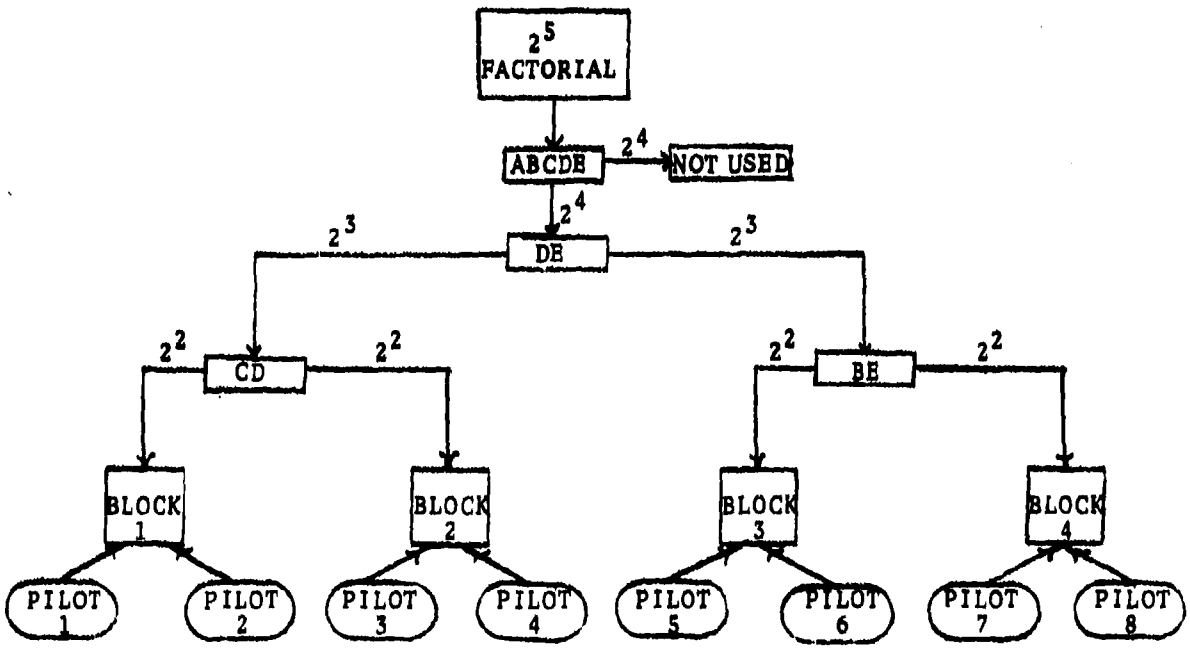


Figure 4.- Experimental design.

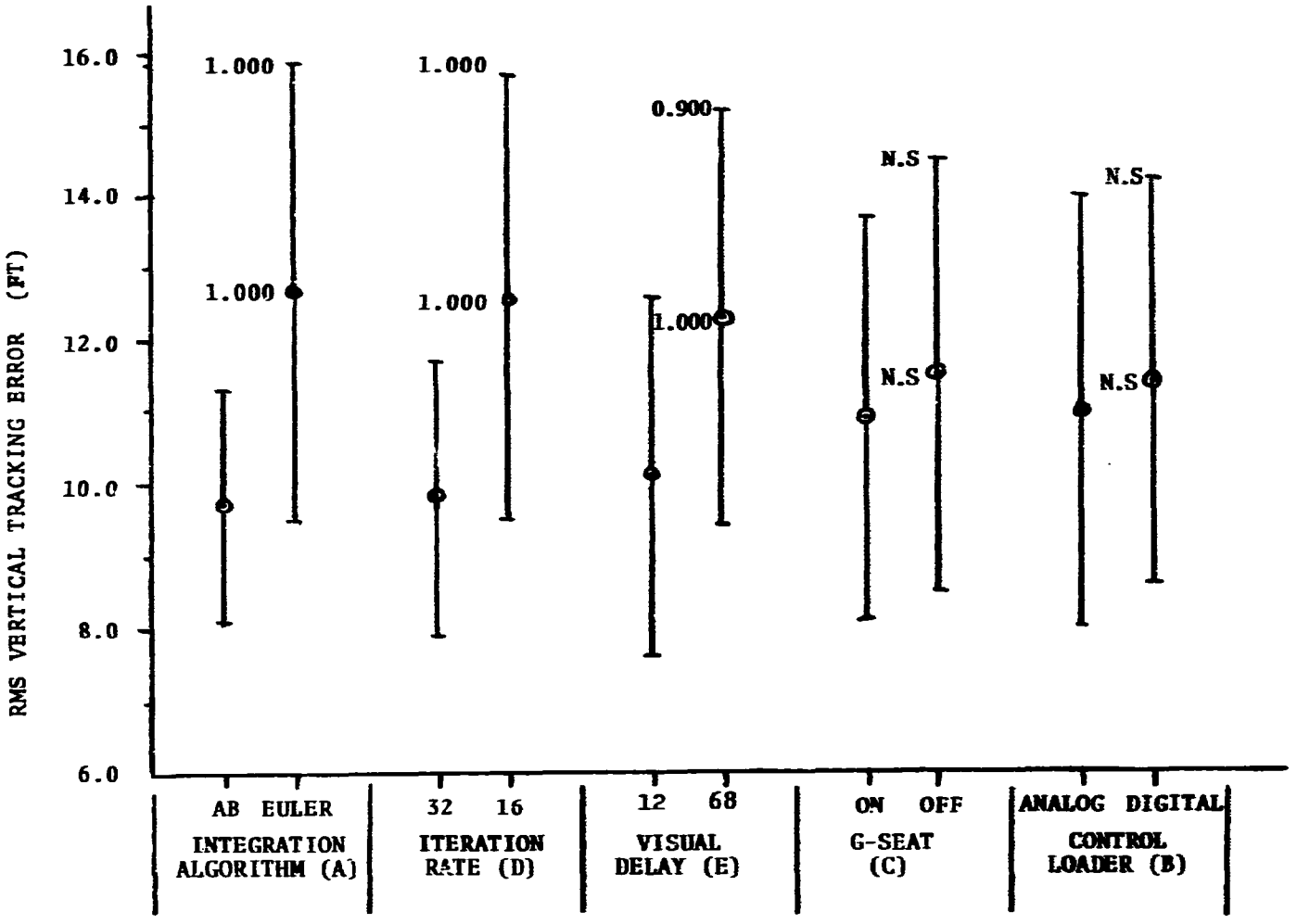


Figure 5.- Tracking error effect for each factor.

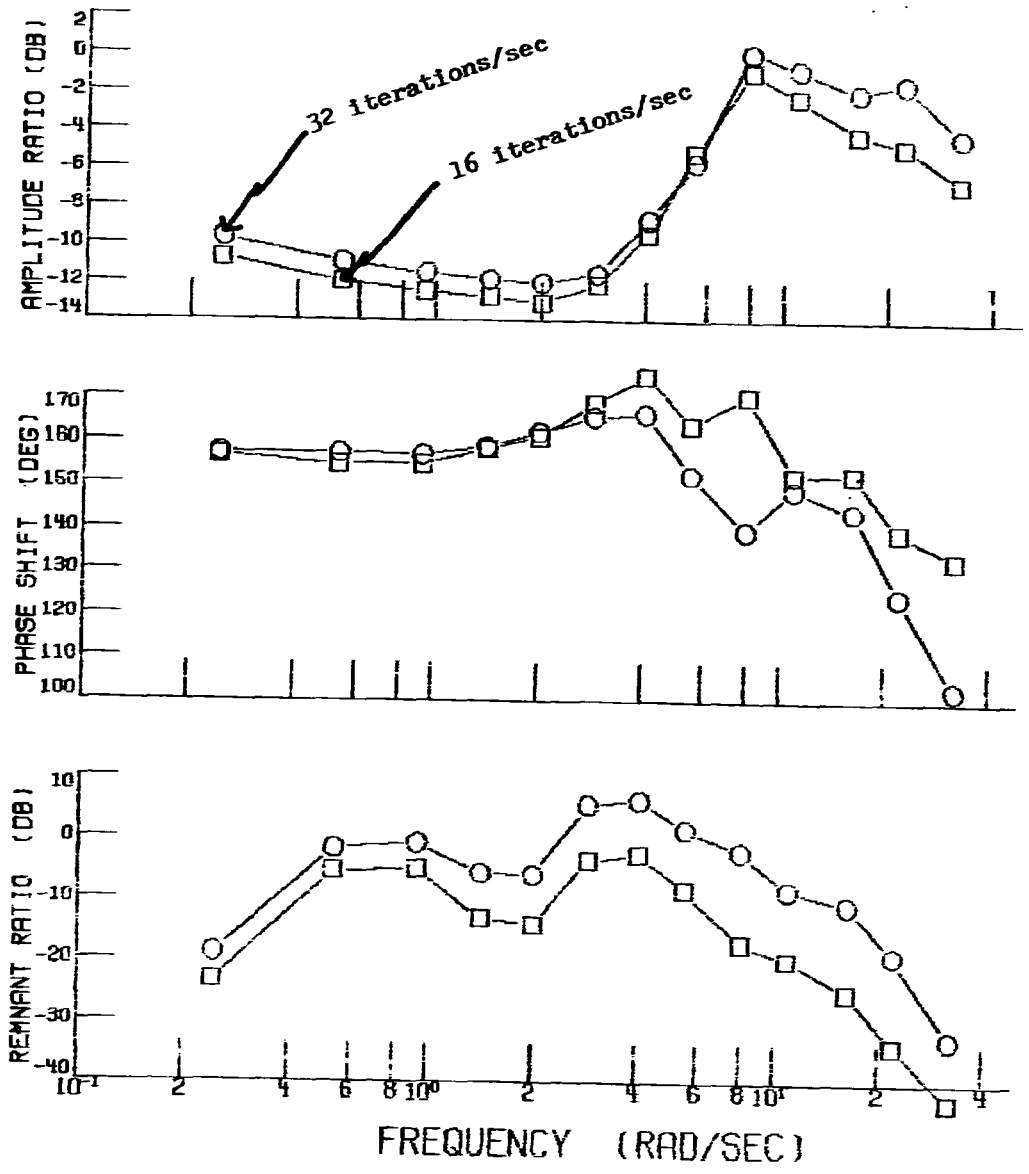


Figure 6.- Pilot describing functions for computer iteration rate.

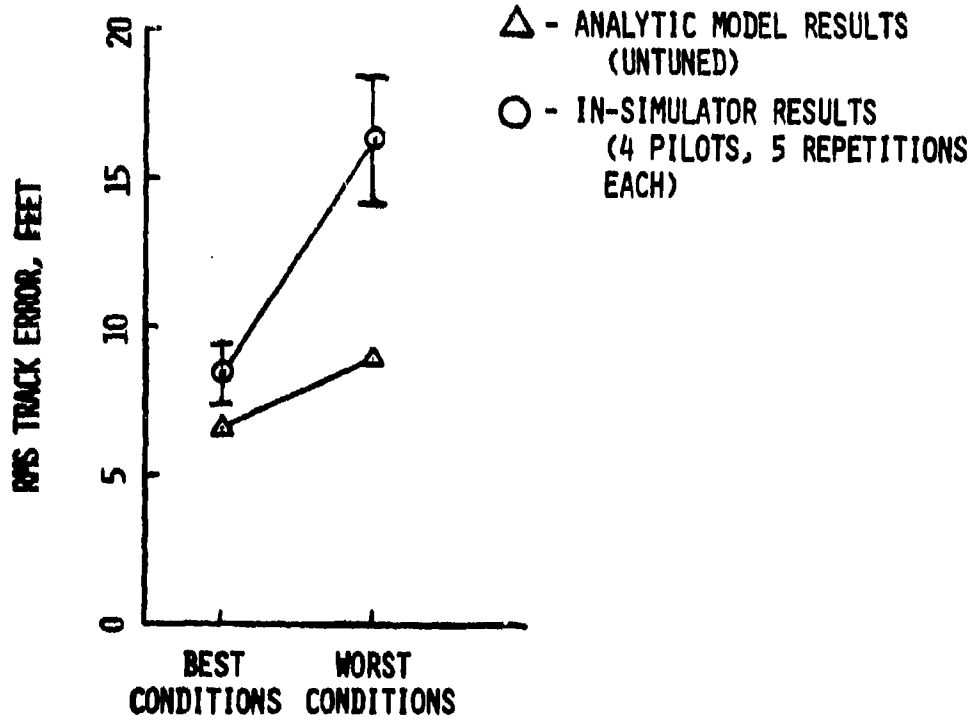


Figure 7.- RMS track error for best vs. worst case.

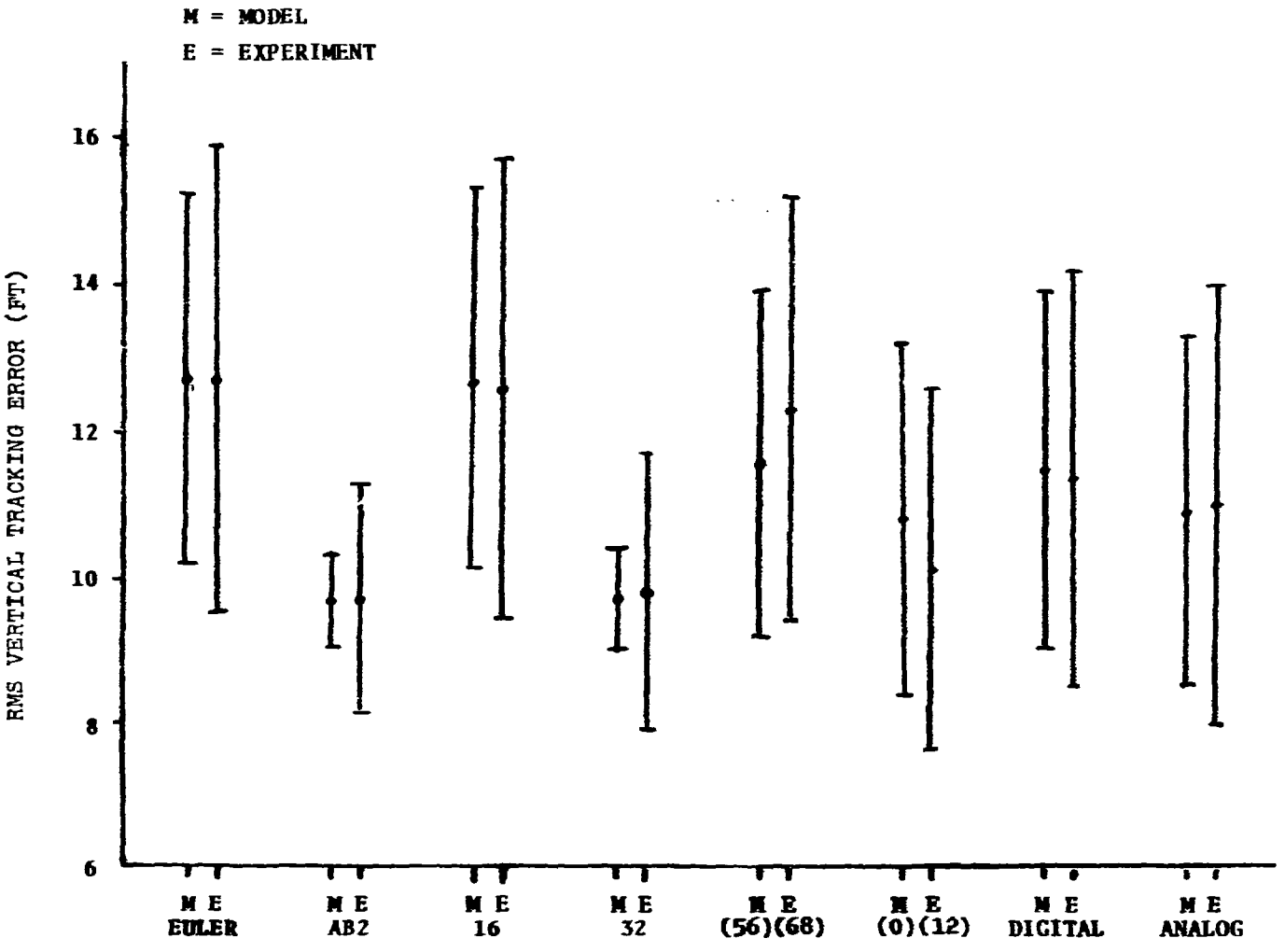


Figure 8.- Comparison of model and experimental tracking error.

EFFECTS OF SIMULATOR DELAYS ON PERFORMANCE AND LEARNING IN A ROLL-AXIS TRACKING TASK

by
William H. Levison, Roy E. Lancraft
Bolt Beranek and Newman Inc.
50 Moulton St.
Cambridge, Ma. 02138

and
Andrew M. Junker
Aerospace Medical Research Laboratory
Wright-Patterson AFB, Ohio 45433

presented at the
Fifteenth Annual Conference on Manual Control
Wright State University, Dayton, Ohio March 20-22, 1979

ABSTRACT

An experimental and analytical study was performed to explore the effects of simulator delays on system performance and operator response behavior during various stages of training on a roll-axis tracking task. Five groups of subjects naive with regard to this task participated. One group trained initially with visual cues only; another group trained with combined, synchronous visual and motion cues; and the remaining groups trained with motion cues delayed with respect to visual cues by amounts of 80, 200 and 300 msec. After initial training, all but the second group were given additional training with synchronous visual and motion cues.

Initial training with motion cues delayed by 80 and 200 msec enhanced learning of the transfer task, as compared to initial training with no motion and with motion delayed by 300 msec. Learning was accompanied by an increase in human operator gain, a decrease in high-frequency phase lag, and a decrease in operator remnant. Analysis with the optimal-control pilot/vehicle model reflected these trends as decreases in time delay, motor time constant, and observation noise/signal ratio.

INTRODUCTION

As part of a long-term study of the effects of motion cues on manual tracking performance, an experimental and analytical study was performed to explore the effects of simulator delays on system performance and operator response behavior during various stages of

training. This study had two major goals: (1) to see how human operators utilize delayed motion cues while learning a control task, and (2) to see how initial training in a task in which motion cues were either non-existent or delayed relative to visual cues would carry over to a situation in which motion cues were synchronous with visual cues.

This paper summarizes the results of the data analysis performed to date. Because only a small subset of the experimental data has been fully analyzed as of the writing of this paper, the results presented herein must be considered tentative, especially with regard to frequency-response and model analysis of performance in early training phases.

DESCRIPTION OF EXPERIMENTS

The experimental task consisted of maintaining a simulated fighter-like aircraft wings level in the presence of random turbulence. Five groups of subjects naive with respect to this task participated. One group trained initially with visual cues only; another group trained with combined, synchronized visual and motion cues; and the remaining groups trained with motion cues delayed with respect to visual cues. All groups were trained to apparent asymptotic mean-squared error performance in their initial tasks. After initial training, all but the second group trained to asymptotic performance with synchronous visual and motion cues.

Experimental Setup

The experimental setup is diagrammed in Figure 1. The subject viewed a display of simulated roll angle, which served as the tracking error, and, by means of a side-mounted control stick, attempted to minimize this error.

Motion cues were provided to the test subjects by the Roll Axis Tracking Simulator (RATS) located at the Aerospace Medical Research Laboratory. The simulator consisted of a single-seat cockpit with a television monitor display and a hand-operated force stick located on the right side. External visual cues were eliminated by a shroud around the cockpit. The axis of rolling rotation was through the human operator's head. The rotating simulator dynamics were identified and simulated on a hybrid computer.

The visual tracking loop contained an analog simulation of vehicle dynamics that included (1) a representation of roll-axis response characteristics of a high-performance fighter aircraft, (2) simulation of the identified dynamics of the rotating

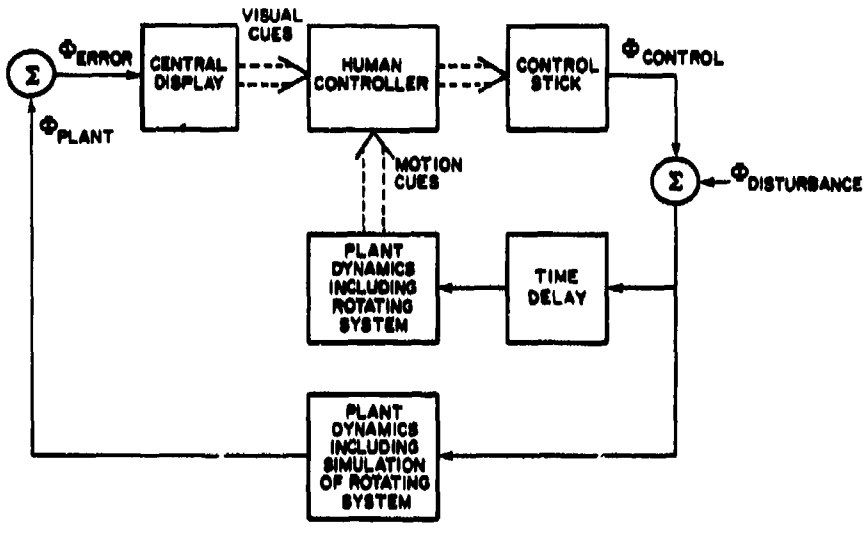


Figure 1. Diagram of the Tracking Task

simulator, and (3) a cumulative transport delay of about 61 msec arising from anti-aliasing filters, processing delays, and delays associated with display generation.

When the tracking task involved combined visual and motion cues, simulator movement was delayed via a digital delay algorithm to provide motion cues that were either synchronous with visual cues (time delay = 0) or were delayed relative to visual cues by amounts of up to 300 msec. The dynamical response properties of the moving-base simulator were augmented with analog dynamics so that when the time delay shown in Figure 1 was set to zero, the motion of the simulator occurred with the same delay and response characteristics as the visual display.

Tracking Display

The negative of the simulated roll angle was displayed as an error to the human operator on a 9-inch diagonal television monitor. The inside-out display consisted of a 1.25 inch long rotating line whose center was superimposed upon a stationary horizontal line. A 0.22 inch perpendicular line at the center of

the rotating line provided upright orientation. The display was centered in azimuth a distance of 26 inches from the subject's eyes. Subjects' sitting heights were such that the display was within 10 degrees of eye level for each subject.

Tracking Task Parameters

The dynamics of the rotating simulator were identified as a low-pass filter having a single root at 19 rad/sec. A second low-pass filter having a single pole at 5 rad/sec was added to these dynamics to approximate aircraft dynamics. With the force stick gain adjusted to produce a steady state roll rate of 14 deg/sec for one pound of force at middle finger height, the controlled dynamics were:

$$\frac{\phi(s)}{U(s)} = \frac{14}{s} \cdot \frac{5}{s+5} \cdot \frac{19}{(s+19)} \tag{1}$$

where $\phi(s)$ is roll angle in degrees and U(s) is control force in pounds.

The disturbance input was constructed from 13 sinusoids whose amplitudes were selected to simulate a random noise process having a power spectral density of the form

$$\Phi_{ii}(\omega) = \left| \frac{K}{(j\omega+2.0)^2} \right|^2 \tag{2}$$

Input amplitude was adjusted to provide an rms disturbance input of 14 deg/sec reflected to the display. In order to prevent subjects from learning the input waveform during the experiments, a random number generator was used to vary phase relationships among component sinusoids from one experimental trial to the next.

Experimental Procedures

Test subjects, all college students, were randomly divided into five groups of four to five subjects each. All subjects were right-handed, in apparently good physical condition, and had no prior tracking experience of the type experienced in this experiment. Five experimental conditions were investigated: (1) visual cues only ("static" tracking), (2) combined synchronous visual and motion cues ("synchronous motion" or "zero delay"), (3) combined visual and motion cues with motion delayed 80 msec with respect to visual cues ("80 msec" condition), (4) motion cues delayed by 200 msec, and (5) motion cues delayed by 300 msec. In all cases, an additional delay of about 61 msec was imposed on both visual and motion cues because of irreducible delays inherent in the simulation procedure.

In order to establish basic capability as a controller, each subject performed two trials of a simulated pitch tracking task using simulated dynamics different from those employed in the subsequent roll-axis task. Each of the five subject groups was then given initial training on one of the five experimental conditions. Training on these tasks continued until asymptotic levels of mean-squared error had been reached. At this point, subjects in the static, 80 msec, 200 msec, and 300 msec groups were transitioned to the synchronous motion case and trained to apparent asymptote on this task.

Each subject completed one session of four experimental trials per day. Each experimental trial lasted 165 seconds. Since learning data were desired, subjects were initially given only the briefest possible familiarization with the simulation; time histories for all subsequent pre- and post-transition experimental trials were recorded for analysis and modeling.

Subjects were instructed to minimize mean-squared error and were shown their error score at the end of each trial. Subjects were made aware of their previous scores as well as those of other subjects in the same group to provide motivation. Between-group performance was kept secret from each group.

EXPERIMENTAL RESULTS

Learning Trends

Pre- and post-transition learning trends for the various subject groups are revealed by changes in mean squared error performance with training as shown in Figure 2.

To obtain the data points shown in this figure, within-subject averages were obtained across the four error scores yielded by each subject in a given session. Across-subject averages were then obtained for all subjects in a given group. To facilitate graphical presentation, average scores -- connected by straight lines -- are shown for every other session prior to transition.

As expected, the introduction of simulator delay degraded asymptotic mean-squared error performance on the initial task. Error scores ranged from about 7.5 deg^2 for synchronous motion to about 40 deg^2 for the 300 msec delay. An MS error score of about 33 deg^2 was obtained by the static group. Thus, motion cues with delays of up to 200 msec benefited asymptotic performance. That is, scores obtained under these conditions were smaller than those achieved by the static group. Motion cues delayed by 300 msec were clearly of no benefit and may have caused the error scores to be larger than they would have been if the subjects had tracked fixed base.

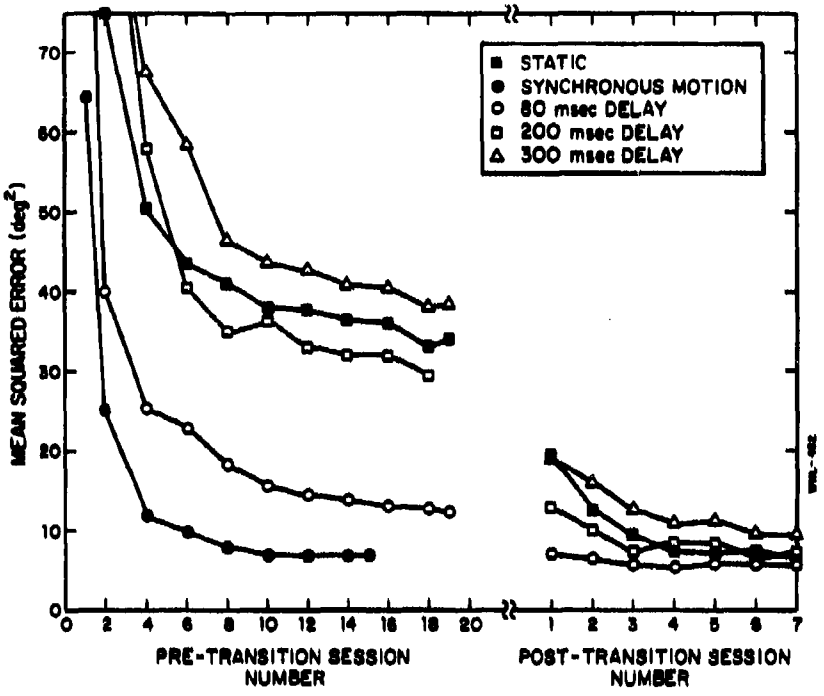


Figure 2. Learning Curves
Average of 4-5 subjects, 4 trials/subject.

Application of the Mann-Whitney non-parametric test for statistical significance [1] revealed significance levels of Ø.Ø5 or less for ordered groups of subjects tracking with combined visual and motion cues. That is, the average MS error achieved by the 8Ø msec group was significantly greater than that obtained by the zero delay group; the 2ØØ msec group had a significantly higher score than the 8Ø msec group, etc. By this measure, asymptotic error scores of the static group were significantly greater than scores yielded by the zero delay and 8Ø msec subject groups. Because of subject-to-subject variability, no statistical significance could be associated with the difference between static performance and either 2ØØ msec or 3ØØ msec performance.

A considerable spread is shown in Figure 2 for average MS tracking error in the first session following transition to the zero-delay motion task. Initial post-transition scores ranged from about 7 for the 8Ø-msec delay task to about 19 deg^2 for the static and 3ØØ-msec delay cases. The Mann-Whitney test revealed statistical significance at the Ø.Ø5 level or less between the 8Ø-

msec, 200-msec, and static conditions. Although Figure 2 shows asymptotic post-transition scores for the 200- and 300-msec cases to be higher than for the static and 80-msec conditions, these differences were not found to be statistically significant.

In order to facilitate comparison of learning rates among the subject groups, learning curves were approximated by an exponential-decay model of the following form:

$$x_n = x_a + (1-F)^n (x_o - x_a) \tag{3}$$

where x_n is the error score after completion of the n^{th} training session, x_a is the asymptotic MS error, x_o is the (theoretical) initial error score before completion of the first training session, and F is the "learning rate", defined as the fraction by which the difference between current error and asymptotic error is reduced from session to session. Thus, if asymptotic error were reached in a single session, F would be unity; if each tracking session reduced tracking error by 30% of the total reduction possible, F would be 0.3; etc.

Model parameters x_a, x_o, and F were found by minimizing the following loss function:

$$E = \left\{ \frac{1}{N} \sum_{i=1}^{N} \left(\frac{m_i - d_i}{\sigma_i} \right)^2 \right\}^{\frac{1}{2}} \tag{4}$$

where N is the number of training sessions considered, m_i is the error score predicted for the i^{th} training session by the above model, d_i is the corresponding "data" (MS error averaged over a subject group), and σ_i is the standard deviation of the subject means within the group. Eight training sessions were considered in computing model parameter values.

The "matching error" E indicates the average number of standard deviations by which model predictions differ from experimental subject means. Note that the "initial" and "asymptotic" error x_o and x_a are theoretical quantities and are not necessarily identified with specific individual measurements. Values for the parameters of the exponential learning model are given in Table 1 for each subject group in pre- and post-transition training.

Although no monotonic relation between motion-cue delay and learning rates was found, Table 1 does show some consistent trends. Specifically, subjects trained initially with the 300-msec motion delay showed collectively the lowest learning rates both before and after transition to the synchronous motion condition. The subject

TABLE 1
PARAMETER VALUES FOR THE EXPONENTIAL LEARNING MODEL

CONDITION	ASYMPTOTIC MS ERROR	INITIAL MS ERROR	LEARNING RATE	MATCHING ERROR
A) Pre-Transition				
No Delay	7.2	37.	.36	.26
80 Msec	13.	64.	.25	.34
200 Msec	32.	181.	.36	.32
300 Msec	37.	97.	.18	.34
Static	35.	104.	.27	.30
B) Post-Transition				
80 Msec	5.7	9.8	.63	.29
200 Msec	7.0	27.	.81	.37
300 Msec	8.2	24.	.37	.13
Static	6.5	33.	.52	.08

populations trained with delays of 200 msec or less tended to reach asymptotic performance as fast or faster than those trained with no motion cues. Thus, initial training with delayed motion cues appeared to enhance the rate of learning of both pre- and post-transition tasks when delayed by 200 msec or less and to impede learning when delayed by 300 msec.

The "learning rates" given in Table 1 provide comparisons of the rates at which subjects reached asymptotic performance on the various task configurations. A more relevant indicator of training effectiveness is perhaps given by initial performance following transition, as one is typically interested in how well the trainee has learned to fly the airplane (represented by the zero-delay

condition) -- not how well he has learned to fly the simulator (represented by the remaining conditions).

Figure 2 shows that initial post-transition MS error scores were substantially lower for the 80- and 200-msec conditions, relative to the 300-msec and static conditions; as noted above, these differences were statistically significant. Thus, in terms of providing training relevant to the post-transfer task, motion simulation with delays of 200 msec or less would appear to enhance learning, compared to static simulation, for the particular tracking task employed in this study.

Frequency Response and Model Analysis

Figure 3 shows the effects of training on frequency-response measures for a single subject in the pre-transition static condition. Response variables are the amplitude ratio and phase shift components of the operator's describing function, and the remnant-related control spectrum (i.e., the portion of the operator's control response that is not linearly correlated with the external forcing function.)* Measures are shown for early and late training.

Training influenced all three frequency-response measures. The operator substantially increased her control gain during the course of training and generated less phase lag at the highest frequency at which reliable measurements could be obtained.** The remnant spectrum was also effected by training: low-frequency remnant decreased and high-frequency remnant increased, as did the frequency at which remnant power was maximal. Measurements obtained at an intermediate level of training (not shown in Figure 3) revealed a consistent trend.

Figures 4 and 5 show that training had similar qualitative effects on amplitude ratio and phase shift for the same subject's post-transition exposure and another subject's initial exposure to the zero-delay motion condition. Effects on remnant control power

---

* In all figures showing frequency response, 0 dB amplitude ratio indicates 1 pound control force per degree roll angle, and 0 dB control power indicates 1 pound^2 control power per rad/sec.

** Estimates of amplitude ratio and phase shift were considered reliable only if input-correlated control power was at least 6 dB greater than remnant-related power at neighboring frequencies. These measurements were not plotted at frequencies for which this criterion was not met. Estimates of remnant-related power were plotted over the entire range of 0.2 to 11 rad/sec.

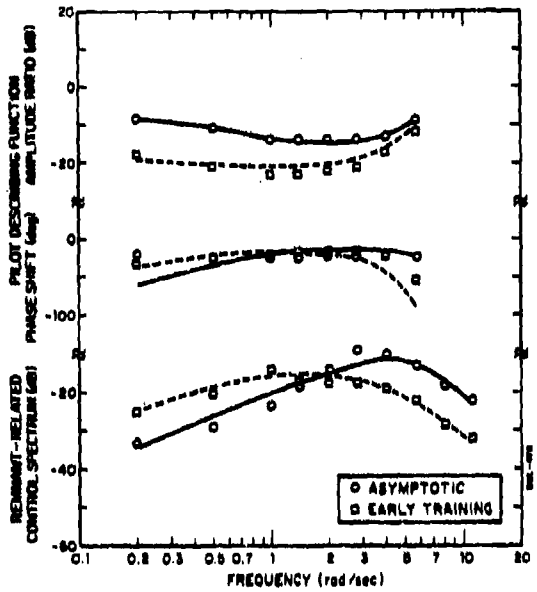


Figure 3. Effects of Training on Pre-Transition Frequency Response, Subject CP, Static Average of 4 trials.

differed from those shown for initial static training in that training with synchronous motion resulted in an overall reduction in remnant power.

Model analysis was performed on selected results with the "optimal-control" model for pilot/vehicle systems. Values for pilot-related model parameters were identified from experimental data as described in a companion paper [2]. The reader is assumed to have a working knowledge of this model.

Analysis was performed on data obtained from subjects trained initially in the static and zero-delay motion tasks; parameter values identified from these results are given in Table 2. Columns labeled "early" refer to data obtained in the first two sessions of the particular training condition, whereas "late" refers to asymptotic tracking performance.

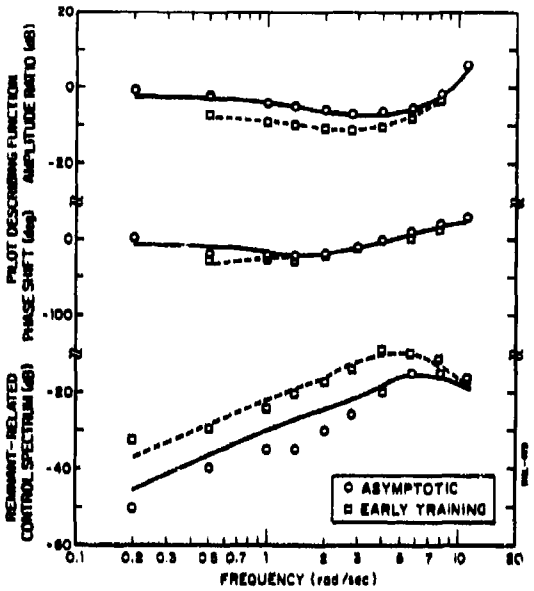


Figure 4. Effects of Training on Post-Transition Frequency Response, Subject CP, Zero Delay Motion

Average of 4 trials.

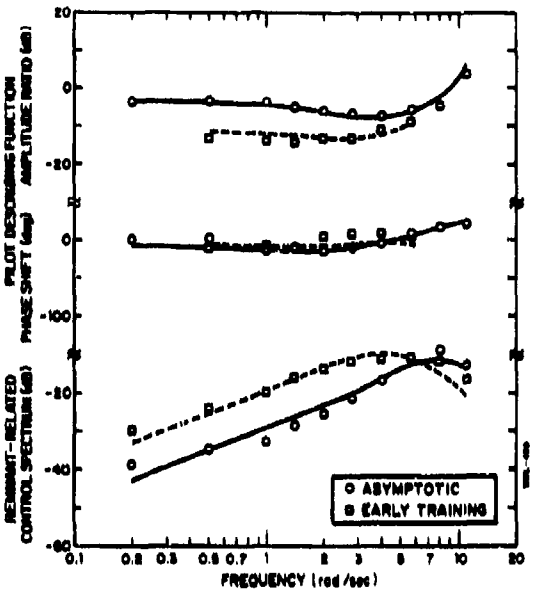


Figure 5. Effects of Training on Frequency Response, Subject DM, Zero Delay Motion

Average pre-transition performance for static tracking early in training was obtained by identifying parameters for each of four subjects, then averaging the identified values across subjects (column 1 of the table). Averaging of the parameters, rather than matching average response behavior, was deemed more appropriate for these early results because of the potential for large individual differences. Because asymptotic performance was less variable across subjects, however, frequency-response measures obtained late in training were averaged across subjects and parameters were identified from the averaged data (column 2). Parameters were also identified from data obtained from individual subjects for synchronous motion tracking -- one subject transitioning from initial static tracking, and one subject trained entirely with synchronous motion.

Model parameters shown in Table 2 include time delay; motor time constant (related inversely to response bandwidth); and

TABLE 2
PILOT-RELATED OPTIMAL CONTROL MODEL PARAMETERS

Model Parameter	Static Pre-Transition		Motion (3) Post-Transition		Motion (3) (No Transition)	
	Early(1)	Late(2)	Early	Late	Early	Late
Time Delay (Seconds)	.28	.19	.16	.15	.28	.19
Motor Time Constant (seconds)	.21	.13	.12	.09	.13	.08
Position Noise/ Signal Ratio (dB)	-7	-22	-21	-28	-17	-31
Rate Noise/ Signal Ratio (dB)	-14	-17	-19	-29	-17	-22
Acceleration Noise/Signal Ratio	---	---	-16	-17	19.5	-23

(1) Average of parameters from 4 subjects

(2) Parameters for average response of 5 subjects

(3) Parameters for 1 subject

observation noise/signal ratios related to perception of error displacement, error rate, and error acceleration. Motor noise parameters were also identified, but as tracking performance was related only weakly to motor noise, these results are omitted from the table and from the following discussion.

When tracking with visual cues only, the subject was assumed to utilize only position and rate information; when tracking with combined visual and motion cues, the subjects were assumed to obtain and utilize displacement, rate, and acceleration

information. Since only one observation noise variance can be uniquely identified for a given perceptual variable, the observation noise quantities shown in Table 2 reflect the combined effects of visual and motion cue utilization.*

Pre-transition training by the static subject group was marked by substantial reductions in time delay, motor time constant, and noise/signal ratio associated with perception of error. Only a modest decrease was observed for rate-related observation noise. Asymptotic values for time delay and noise/signal ratios were on the order of those observed in early laboratory studies with wide-band tracking dynamics; the average motor time constant, however, was about 30% greater than found in these earlier studies.

Transition to zero-delay motion was accompanied by further reductions in motor time constant and noise/signal ratios, but no further reduction in time delay. Asymptotic motor time constant for this condition was consistent with earlier results, but asymptotic noise/signal ratios for displacement and rate were close to -30 db -- about 10 dB lower than levels previously identified from tracking tasks involving stable plant dynamics.

Analysis of the data obtained from one subject in the zero-delay motion group also revealed substantial reductions in all pilot-related model parameters (except for acceleration-related noise/signal ratio, which was relatively low early in training). Asymptotic values for time delay and motor time constant were similar to those observed in previous studies, whereas noise/signal ratios were, on the average, lower than expected for a stable tracking task.

In summary, the following trends are apparent from the data analyzed so far:

1. Initial training on the static task was accompanied by substantial reductions in time delay, motor time constant, and observation noise related to perception of displacement.

2. Additional training following transition to the task with synchronized visual and motion cues allowed an immediate drop in time delay, which remained stable with further training, plus progressive decreases in motor time constant and noise/signal ratio.

---

* One could attempt to distinguish visual from motion cue effects by making certain constraining assumptions, such as assuming that attention is shared between motion and visual cues according to some model for attention-sharing [7]. However, no such attempt was made here.

3. Initial training with synchronous visual and motion cues revealed progressive decreases in all pilot-related model parameters.

Some of these trends confirmed prior expectations, while others did not. Additional model analysis was undertaken to explain these results, as described below.

Discussion of Results

It was anticipated that training would be accompanied by a progressive increase in the operator's amplitude ratio, a decrease in high-frequency phase lag, and a reduction in remnant. In general, any reduction in what we consider to be basic information-processing limitations (time delay, motor time constant, noise/signal ratio) would be accompanied by these changes in frequency response.

Although progressive reductions in all three parameters were expected to some extent, the trends exhibited specifically by the noise/signal ratio and motor time constant deviated from prior expectations.

Let us first consider the observation noise/signal ratio. This parameter provides the major source of "pilot remnant" in the optimal-control pilot/vehicle model. Therefore, any type of pilot activity that cannot be accounted for by a fixed linear response strategy contributes to this noise parameter. Such activity may include nonlinear behavior, injected noise, and variations of the linear response strategy during the course of an experimental trial.

We expected that the operator's attempt to discover the appropriate response behavior would cause him, during the early training phase, to vary his control strategy during the course of individual experimental trials. Such behavior would be manifested by large noise/signal ratios. Moreover, we anticipated that subjects would have difficulty utilizing velocity information early in training and that observation noise associated with rate perception would be especially large.

Trends shown in Table 2 for static tracking are the opposite of what was anticipated. Rate-related observation noise early in training was within 3 dB of that which matched asymptotic static tracking, whereas a very large noise/signal ratio was initially found for position-related noise.

One should perhaps exercise a good deal of caution, however, when interpreting observation noise levels identified for untrained subjects. While the relation between noise/signal ratios and cue utilization appears to be justified for the highly-trained

subject [3,4], other factors (such as non-optimal response behavior) may influence these pilot-related model parameters early in training.

The relatively low noise/signal ratios achieved after training with combined, synchronous visual and motion cues were also surprising, because (1) such low noise levels (on the order of -26 dB) had been observed in earlier studies primarily when vehicle dynamics were unstable [5,6], and (2) a previous experiment using a similar tracking task yielded noise/signal ratios on the order of -20 dB [7].

Sensitivity analysis provides at least a partial explanation for the low noise/signal ratios. If we set other pilot parameters equal to values appropriate to the asymptotic motion results, reducing the noise/signal ratio from -20 to -26 dB reduces the predicted mean-squared error to .37 of its value for the -20 dB noise level. This sensitivity is on the order of that found in the experiments with unstable dynamics. On the other hand, a similar change in noise/signal ratio for the static case (where other parameters are readjusted to provide the best match for that condition) reduces the predicted error to 0.57 of the initial value. Thus, sensitivity analysis provides a rationale for the different asymptotic noise/signal ratios observed in this experiment.

Differences in noise/signal ratios found in different motion-base experiments may also be explained via sensitivity analysis. In a recent experiment employing the same task dynamics and inputs, subjects were trained to minimize a weighted sum of mean-squared error and mean-squared plant acceleration, rather than simple mean-squared error [7]. Reduction of noise/signal ratio from -20 to -26 dB reduced the predicted total "cost" to .66 of its initial value. Thus, the relatively lower sensitivity of performance to observation noise may have failed to motivate the subjects in the earlier experiment to lower their noise/signal ratios beyond the -20 dB normally associated with laboratory tracking tasks.

Combined pilot/vehicle frequency response behavior for selected subjects was analyzed to determine whether or not the relatively large motor time constants reflected attempts by the subjects to adopt conservative response strategies that would provide greater margins of system stability than that associated with the predicted optimal response behavior. Specifically, gain crossover frequencies and phase margins were derived from the frequency response measurements, where "gain crossover" is the frequency at which the combined pilot/vehicle system responds with unity amplitude (i.e., 0 dB), and "phase margin" is the phase shift of the total system at gain crossover plus 180 degrees. Since, in general, the system would be unstable if the phase shift were less than -180 degrees at gain crossover, the phase margin may be considered an indicator of relative system stability.

Table 3 shows that phase margin was reduced considerably from that obtained in early training to that achieved late in pre-transition training. Asymptotic phase margins were about 25 degrees for static tracking and 21 degrees for the subject trained with motion cues, whereas phase margins early in training were over 40 degrees. Thus, it appears that the subjects adopted large motor time constants (i.e., reduced their response bandwidths) in early training to achieve greater system stability, perhaps to compensate for their lack of knowledge concerning vehicle response behavior.

TABLE 3
COMBINED PILOT/VEHICLE RESPONSE PARAMETERS

Task	Subject	Phase Margin (degrees)	Gain Crossover (rad/sec)
Early Static	CP	52	1.1
Asymptotic Static	CP	25	2.7
Asymptotic Static	Average	26	3.0
Model Prediction, Static	---	25	3.5
Early Motion, Post-Transition	CP	22	3.9
Asymptotic Motion, Post-Transition	CP	19	5.2
Early Motion, Pre-Transition	DM	45	3.3
Asymptotic Motion, Pre-Transition	DM	21	5.1

The larger-than-expected motor time constant found for asymptotic static tracking does not appear to have increased system stability, however. The phase margin predicted by the model using values for pilot parameters consistent with earlier results was about 25 degrees for static tracking -- about the same as that found experimentally. Predicted gain crossover frequency was greater than experimental values, however (3.5 rad/sec compared with 3.0 for the average of five subjects and 2.7 for subject CP), and predicted MS error was correspondingly lower than that found

experimentally. It is therefore not clear what benefit was derived by maintaining a motor time constant above the nominal 0.1 seconds.

The average asymptotic motor time constant of 0.13 sec for static tracking does not appear to reflect an inherent bandwidth limitation of this particular subject group. For example, a subject having a static motor time constant of 0.18 seconds provided data that was matched with a time constant of 0.1 seconds when she subsequently trained on the post-transition task.

Motor time constants greater than 0.1 seconds have been observed in earlier experiments involving relatively low-bandwidth plant dynamics [4,8]. Although plant bandwidth in this experiment was not especially low, the bandwidth limitations of the simulator (a single pole at 19 rad/sec) plus the simulation delay of about 60 msec, when combined with the basic plant dynamics, served to induce larger high-frequency phase lags than have been found in many of the early experimental studies with simple tracking dynamics. When viewed in this context, the experimental results are not inconsistent with previous findings.

The apparent decrease in pilot response bandwidth with low system bandwidth (or large system phase lag) remains to be explained. As we have just shown, model analysis does not support the hypothesis that such behavior represents an optimal strategy, since superior performance could be presumably achieved if the subject were to operate at higher bandwidths.

The inability of the subject to develop an adequate internal model of the task environment in certain situations may be in part responsible for increased motor time constant. That is, increasing the time constant may, in fact, represent an optimal strategy on the part of the subject when forced to respond in situations of significant uncertainty. Alternatively, the control system may not provide enough information for the subject to generate a high response bandwidth.

SUMMARY

Analysis completed to date yields the following tentative conclusions:

1. For the specific tracking task explored in this study, asymptotic pre-transition tracking performance was influenced by simulator delays. Mean-squared error significantly increased for increasing delay. Delays of 200 msec or less led to lower scores than did static (fixed-base) tracking; whereas scores obtained with motion cues delayed by 300 msec were, on the average, greater than static scores.

2. Initial training with motion cuess delayed by 80 and 200 msec enhanced learning of the transfer task, as compared with initial training with no motion or with motion delayed by 300 msec.

3. Pre- and post-transition learning was accompanied by increasing operator amplitude ratio and decreasing phase lag; post-transition learning was also accompanied by decreasing remnant.

4. The following effects of training on identified pilot-related model parameters were observed:

   a. Initial training on the static task was accompanied by substantial reductions in time delay, motor time constant, and observation noise/signal ratio.

   b. Additional training following transition to the task with synchronized visual and motion cues allowed an immediate drop in time delay, which remained stable with further training, plus progressive decreases in motor time constant and noise/signal ratios.

   c. Initial training with synchronous visual and motion cues was accompanied by progressive decreases in all pilot-related model parameters.

5. The motor time constant found for asymptotic static tracking was greater than generally found in the past for simple tracking dynamics, whereas the noise/signal ratio for asymptotic zero-delay motion tracking was lower than found in recent studies of motion cues.

6. Differences in noise/signal levels found in this and prior studies can be explained on the basis of differential sensitivity of system performance to noise/signal ratio.

7. The cause for the elevated motor time constant is not clear. Further study is suggested to explore the hypothesis that larger motor time constants, especially early in training, are related to uncertainties on the part of the human operator.

As of the date of this paper, model analysis of the 80-, 200- and 300-msec results has not yet been performed. Analysis of the data obtained in this study is expected to continue under a current study for the Air Force.*

---

* AFOSR Contract No. F49620-78-C-0073

REFERENCES

1. Siegel, Sidney, Nonparametric Statistics for the Behavioral Sciences. New York: McGraw-Hill Book Company, Inc., 1956.

2. Lancraft, R. E. and D. L. Kleinman, "On the Identification of Parameters in the Optimal Control Model," Proceedings of the Fifteenth Annual Conferences on Manual Control, March 20-22, 1979, Dayton, Ohio.

3. Levison, W. H., S. Baron, and D. L. Kleinman, "A Model for Human Controller Remnant," IEEE Trans. on Man-Machine Systems, Vol. MMS-10, No. 4, December 1969.

4. Levison, W. H., "The Effects of Display Gain and Signal Bandwidth on Human Controller Remnant," AMRL-TR-70-93, Wright-Patterson Air Force Base, Ohio, March 1971.

5. Levison, W. H., J. I. Elkind and J. L. Ward, "Studies of Multivariable Manual Control Theory," NASA CR-1753, June 1971.

6. Zacharias, G. L. and W. H. Levison, "A Performance Analyzer for Identifying Changes in Human Operator Tracking Strategies," BBN Report No. 3910, Bolt Beranek and Newman Inc., Cambridge, Mass., August 1978.

7. Levison, W. H. and A. M. Junker, "A Model for the Pilot's Use of Motion Cues in Roll-Axis Tracking Tasks," AMRL-TR-77-40, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, June 1977.

8. Levison, W. H., "A Model for the Pilot's Use of Roll-Axis Motion Cues in Steady-State Tracking Tasks," BBN Report No. 3808, Bolt Beranek and Newman, Inc., Cambridge, Mass., May 1978.

ACKNOWLEDGEMENTS

The experimental study was a cooperative effort between AMRL/ME and AFHRL/AS and was performed at AMRL. Primary data reduction was performed at AMRL, and model analysis and interpretation of results was provided by Bolt Beranek and Newman under the following contracts: (1) Systems Research Laboratory, Purchase Order No. 23676 in support of Project No. 6872-12-31, Contract No. F33615-76- C-5001, and (2) AFOSR Contract No. F49620-78-C-0073.

A PERFORMANCE ANALYZER FOR IDENTIFYING CHANGES IN HUMAN OPERATOR TRACKING STRATEGIES

Greg L. Zacharias and William H. Levison
Bolt Beranek and Newman Inc.
50 Moulton Street
Cambridge, MA

ABSTRACT

A study was undertaken to develop, implement, and validate a compensatory tracking task for use in assessing visual countermeasures. The human operator optimal control model was used to design the task, and a minicomputer-based system was developed for implementation and experimental validation. Two simulated stressors were studied, and it was shown how data trends could be used to identify particular stressor effects. Quantitative model analysis provided an efficient means of compressing the tracking data and of identifying stressor effects on the operator's information processing capabilities.

1. INTRODUCTION

This paper summarizes the results of a tracking task development effort conducted by Bolt Beranek and Newman Inc. (BBN) for the Aerospace Medical Research Laboratory (AMRL/HEA); detailed study results are documented by Zacharias and Levison (1978).

The objective of the study was to develop, implement, and validate a compensatory tracking task having a wide applicability and sensitivity to a number of operator stressors present in the operator's environment. The task was to be designed for use with the type of stressors expected to be utilized in a program investigating the effects of visual countermeasures, thus serving as a preliminary screening task for countermeasures assessment. Task development was to be based on current human operator modelling theory, to allow for the isolation of the essential elements of operator performance characterizing tracking under stressed conditions.

Our tracking task development effort had two specific objectives. First, we wished to design a task in which operator performance was a sensitive function of the stress imposed on the operator. Second, we wished to define a set of performance metrics which would provide us the capability for differentiating among qualitatively different stressors. These objectives thus required a task having not only general sensitivity to imposed stressors, but a differential sensitivity to different types of stressors.

At the outset, we recognized that task sensitivity is a strong function of the difficulty level associated with the task dynamics. If the dynamics are

particularly easy to control, then performance tends to be an insensitive function of task stress; conversely, dynamics which are controlled only with great difficulty are associated with performance changes which are highly sensitive to task stress.

A particular example which demonstrates this is the study of hypoxia on tracking, conducted by Replogle et at (1970). Subjects were required to control two sets of dynamics: one a third-order plant typical of aircraft pitch dynamics, and the other a time-varying first-order instability. The results showed that hypoxia produced a significant performance degradation in the more difficult unstable tracking task. but had no measurable effect in the less difficult pitch-tracking task.

The implication is that the task dynamics should be chosen to make the task fairly difficult, if stress effects are to show up. The first-order unstable dynamics of the "critical tracking task," developed and described by Jex et al (1966), and used by Replogle et al (1970), provide this required potential for task difficulty and performance sensitivity.

A drawback of the critical task, however, is the fact that only a single measurement is produced: the maximum value of the unstable plant pole controllable by the subject. Changes in this value, or its inverse, the operator's effective time delay (Jex et al (1966)), might be simply correlated with changes in operator stress, but such changes would not serve as the basis for identifying <u>specific</u> stressor effects. This is because of the fact that <u>any</u> stressor which served to degrade the operator's information processing ability (such as his perceptual threshold), would be interpreted as an increase in operator time delay. Thus, although the critical task provides the required sensitivity, it does not provide the capability for differentiating among qualitatively different stressors.

This deficiency can be remedied, however, by utilizing unstable dynamics with a <u>fixed</u> instability, in conjunction with an input disturbance signal which acts to continuously disturb the system. This "subcritical tracking task" (Jex et al (1966)), allows for an efficient steady-state analysis of operator performance, with an abundance of performance metrics, which provide the potential for differential sensitivity to different stressors. We thus centered our attention on the subcritical task; the next section discusses our efforts in this area.

2. TRACKING TASK DESIGN

Design of the tracking task was conducted by simulating tracking performance under a variety of conditions, using the optimal control model (OCM) of the human operator (see Kleinman et al (1970) for a description of the model). This allowed us to rapidly evaluate different task designs, and examine specific task sensitivity to a number of simulated stressors.

Model treatment of stressor effects was predicated on the notion that a stressor serves to degrade the operator's basic information processing capabilities, and that a particular stressor acts on specific capabilities, depending on the nature of the stressor. In terms of the OCM, a stressor would affect the parameters associated with information-processing ability: the observation noise/signal ratio, time delay, motor time constant, etc. Further, the effect of a specific stressor (such as target camouflage), would be reflected as a change in a specific model parameter (such as observation noise), allowing for a direct categorization of stressors in terms of their effects on specific model parameters associated with information processing.

The basic tracking task we chose to simulate is shown in Figure 1. This compensatory single-axis task utilizes the fixed first-order unstable dynamics of the subcritical task, in conjunction with an input disturbance having a first-order power spectral density function, and an RMS value of unity.

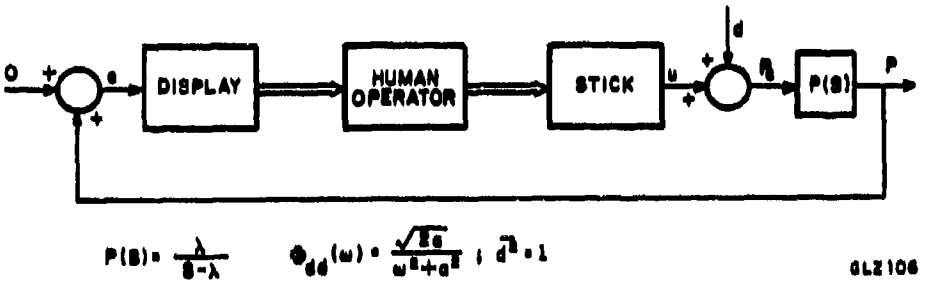


Figure 1: Tracking Task Block Diagram

Our initial design objective was to specify values for the plant pole λ and noise break frequency a. In our simulations, the human operator was replaced by the OCM, and varying stress levels were simulated by varying the model's observation noise/signal ratio. All other model parameters were fixed at "nominal" values consistent with levels associated with previous research: time delay = 200 msec, motor time constant = 100 msec, and motor noise/signal ratio negligibly small. By varying the observation noise about a nominal -23 dB value, we were then able to generate sensitivity curves relating tracking error to observation noise, for a specific task parameter set (λ,a). Several of these simulations (described by Zacharias and Levison (1978)) led us to the conclusion that a moderate instability (λ = 2 rad/sec) combined with a low bandwidth noise (a = 0.5 rad/sec) provided the best compromise between stressor insensitivity and task uncontrollability.

With these task parameter values set, we then proceeded to investigate sensitivity of other task metrics to other model parameter variations. Performance sensitivity to four parameters was explored: time delay, motor time constant, observation noise/signal ratio, and motor noise/signal ratio. The operator was assumed to use both position and rate information, and separate observation noise processes ("position noise" and "rate noise") were associated with these quantities. The corresponding noise/signal ratios were varied jointly to simulate a change in the overall observation noise ratio, and varied individually to represent selective decrements in the ability to obtain and use these informational quantities.

Figure 2 shows the dependence of both tracking error and stick rate variance on changes in the various model parameters. In each plot, the left-most arrow indicates the assumed nominal value for that particular parameter; the next arrow indicates the parameter value associated with a doubling of the nominal score; the right-most arrow indicates a quintupling value. These plots show that, for example, a change in observation noise from -23 dB to -20 dB cannot be differentiated from a change in time delay from 0.2 sec to 0.27 sec when only tracking error is measured, since both parameter changes result in an identical doubling of the score.

At least one of the parameters, however, can be differentiated from the others by considering measures <u>other</u> than tracking score. Figure 2 shows that all of the stick rate scores increase with increasing tracking errors except in the case of the motor time constant. Here stick rate actually decreases, reflecting the lower bandwidth of the human operator. Thus, confronted with a measured increase in tracking score, an experimenter could decide whether or note there has been a change in the motor time constant by simply examining the associated change in the stick rate score.

If it were determined that a change in the motor time constant were <u>not</u> responsible for the tracking error increase, then the experimenter would be faced with differentiating among the remaining parameters. Figure 3 shows the dependence of the normalized remnant spectrum, obtained by reflecting the stochastic portion of the operator's control response to an equivalent perceptual disturbance on error rate, normalized with respect to mean-squared error rate (Levison et al (1969)). Remnant curves are parameterized against those model parameter values associated with the nominal case and with the doubled and quintupled error scores of Figure 2.

A change in each of the model parameters results in a distinctive change in the remnant spectrum. An increase in overall observation noise results in an increase in the remnant level at all frequencies. On the other hand, an increase in position noise selectively increases low-frequency remnant, whereas an increase in rate noise selectively increases remnant at high frequencies. An increase in motor noise also increases high-frequency remnant, but in a different fashion. An increase in time delay results in little remnant change.

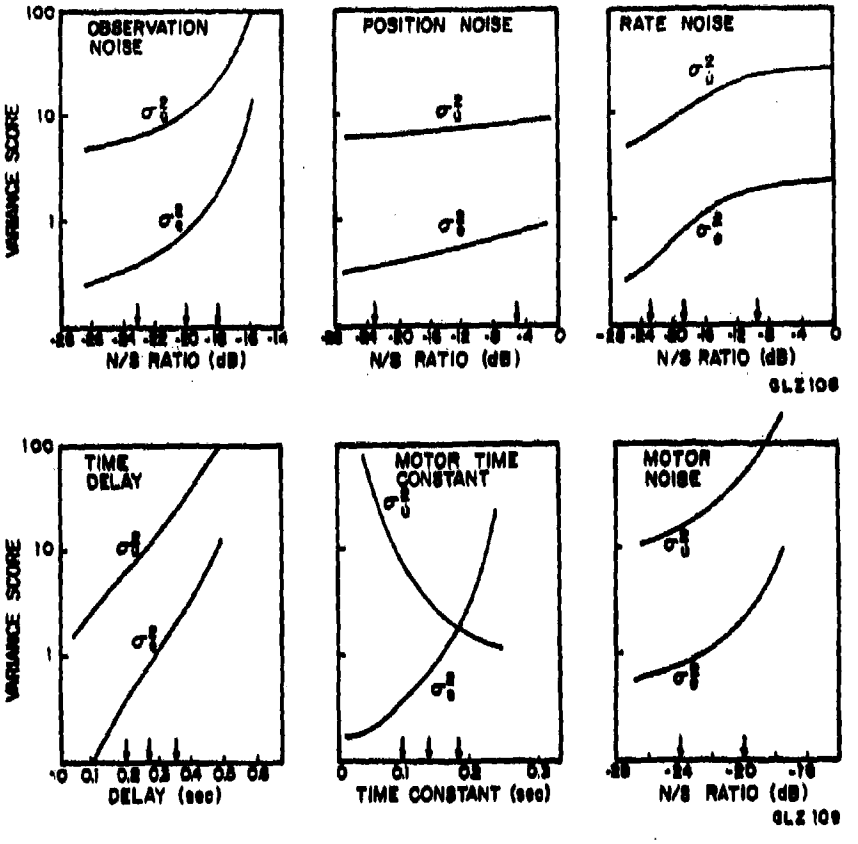


Figure 2. Tracking Score and Stick Rate Sensitivity.

Thus, changes in remnant behavior, in conjunction with relative changes in error and stick rate scores, could allow one to identify the information-processing parameter responsible for the change in tracking performance.

A partial cross-check can be provided by examining changes in the operator's describing function, shown in Figure 4. Changes in position noise have little effect here, and thus the describing function behavior could serve to isolate the position noise parameter from the others. In addition, an increase in the time delay results in a distinctive change in the resonance peak (with respect to both position and sharpness) and in the behavior near resonance; these characteristics could serve to distinguish a change in time delay from changes in the motor time constant and noise parameters.

These results demonstrate that a change in a single information-processing model parameter results in a constellation of changes in the several performance metrics considered. If a stressor serves to directly affect an operator's information processing capabilities, then the distinctive differences

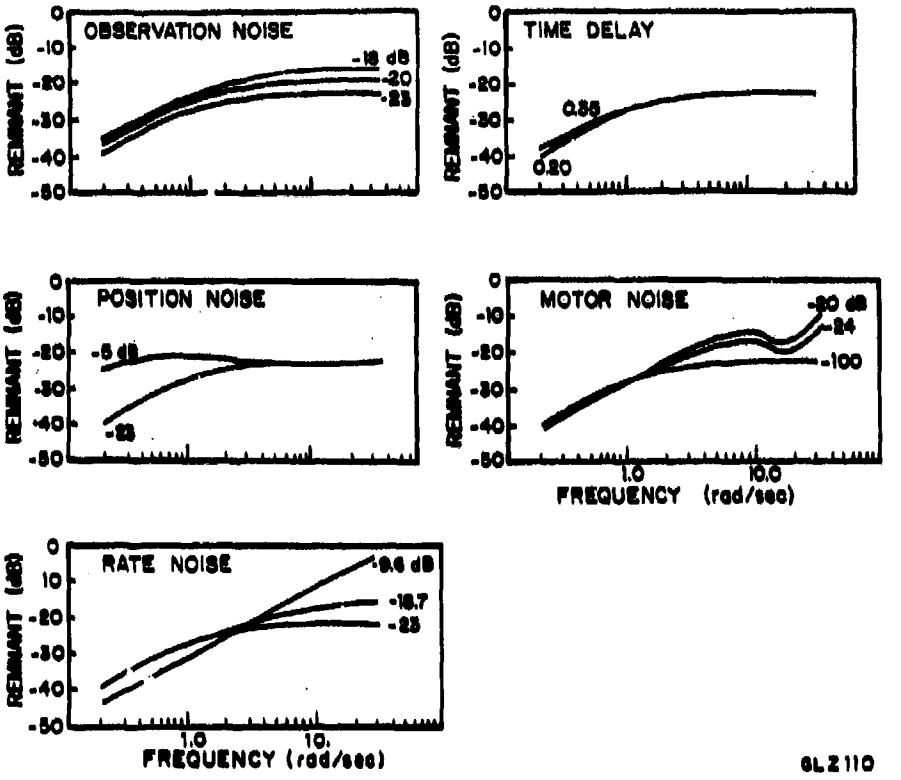


Figure 3. Normalized Remnant Sensitivity

demonstrated here could serve as a basis for identifying the nature of the stressor, in terms of its impact on a specific information processing capability. The task of Figure 1 was chosen as the basic design for our experimental validation effort.

3. EXPERIMENT DESCRIPTION

There were three objectives to our experimental study: a) determine the amount of training needed to ensure consistent operator tracking performance on the chosen task; b) validate the sensitivity of the task to a change in the operator environment; and c) validate the differential sensitivity of the task and its utility in isolating different stressor effects.

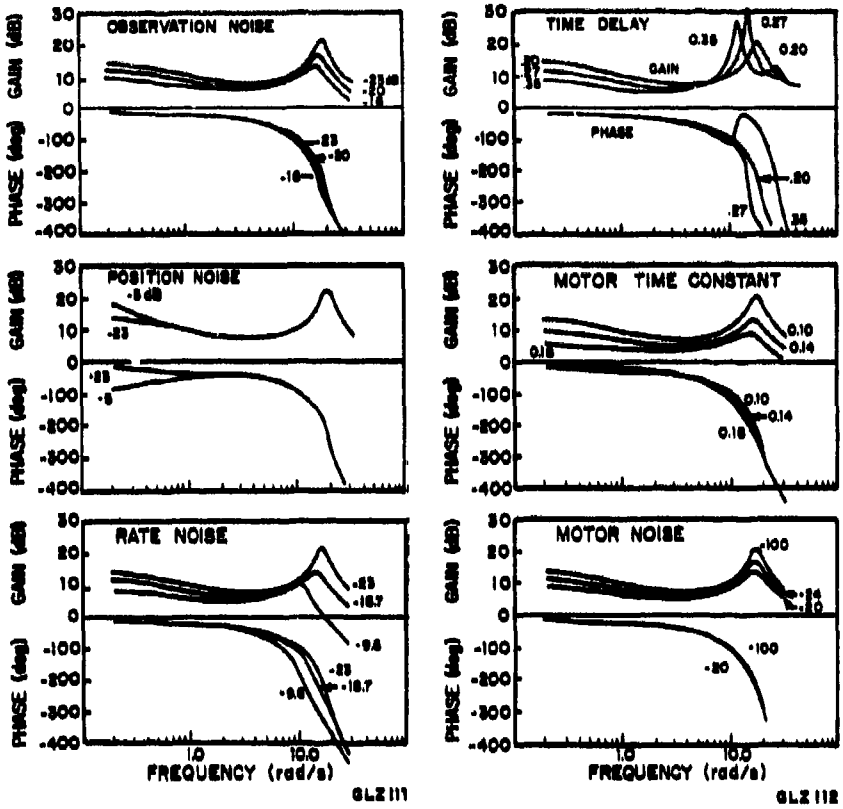


Figure 4. Describing Function Sensitivity

3.1 Task Description

A block diagram of the tracking task has already been illustrated in Figure 1. The subject was provided a compensatory error display and force control stick for stabilizing the plant and compensating against the input disturbance signal.

Tracking error was displayed on a 10 cm by 6 cm oscilloscope face, with 1 cm horizontal and vertical grid lines providing a stationary reference. The outside-in display consisted of a vertical bar 2.6 cm long, deflected proportionally to tracking error. The display was centered in azimuth approximately 60 cm from the subject, so that one centimeter of display displacement resulted in approximately one degree of angular displacement. Subjects' sitting heights were such that the display elevation angle was within 10 degrees of eye level.

Subjects used a left-right force control stick with an arm rest for the subject's arm and hand. Since only 5 cm of the stick were available for

grasping, subjects used a combination of finger and thumb grips for manipulating the stick. Stick stiffness was such that one pound of force applied to the top end of the stick resulted in approximately 1 mm of lateral stick displacement. Stick gain was chosen so that 1 pound of force resulted in a 5 cm plant command, a level which provided adequate controllability without being overly sensitive.

The input disturbance signal was constructed from 13 sinusoids. Their amplitudes were selected so that the input signal power spectral density (PSD) approximated the continuous PSD function given in Figure 1; the break frequency was set at 0.5 rad/sec. RMS input signal level was set at 1 cm.

To prevent subjects from learning the input waveform during the course of the experiment, a random number generator was used to vary the phase relationships of the input signal sinusoids from one tracking run to the next. The input was thus "random-appearing" to the subjects.

The plant dynamics consisted of the first-order instability of Figure 1, cascaded with a simple dead-time associated with the digital implementation of the tracking loop (see below). Plant gain was set at unity, and the instability pole was set to 2 rad/sec. The dead-time was calculated from an input-output system measurement, and found to be 32 msec.

The tracking loop was implemented on a digital computer, with the display and stick providing the analog interfaces to the human operator. The programmable digital system, developed specifically for human operator tracking experiments, is described in detail by Zacharias and Levison (1978). In brief, the system provides the investigator with a simplified means of defining the desired tracking task, automated calibration of the interface hardware, real-time control of the tracking experiment, storage of the tracking time histories, and post-experimental analysis of tracking performance.

To investigate the sensitivity of the tracking task to changes in the task environment, two simulated stressors were used in the experiment. One, the "diffused" condition, simulated "visual noise" imposed on the tracking display, and consisted of a diffused display error bar. By inserting a translucent sheet between the display grid lines and the face of the display, the vertical bar could be diffused over a width of approximately one cm, without affecting the display reference provided by the grid lines. The second, the "delayed" condition, simulated a stressor which acts to increase the operator's central processing time (e.g., via confusion, ambiguity, etc.), and was implemented by adding a time delay to the plant dynamics. A specific delay value of 40 msec was chosen, and was cascaded with the system's 32 msec dead-time mentioned above.

3.2 Experimental Protocol

Six male college students between the ages of 18 and 29 years of age were used as subjects. All were right-handed and had normal visual acuity, although some wore corrective lenses.

A tracking run lasted for 185 seconds. Three such tracking runs, interspersed with one minute rest intervals, comprised a block of runs. Four such blocks, interspersed with 15-minute rest intervals outside of the subject room, comprised a normal day's tracking session for each subject. The overall experiment was conducted over a month's time.

The experiment was conducted in three phases, consisting of an initial training phase, an intermediate training phase on the stressed conditions, and a control phase on the three experimental conditions.

Initial Training Phase

Each subject was instructed as to the task dynamics involved, and the objective of minimizing tracking score. The subject was then allowed an initial practice session to gain familiarity with the task dynamics and the input signal characteristics.

Tracking score was reported to the subject at the end of each run, and the subject was encouraged to modify his hand grip, gaze, and tracking strategy in any manner he felt which would help minimize his tracking score. Each subject was also allowed to view a plot of his run-to-run score history, to assess his progress in learning the task. To maintain motivation, subjects were also made aware of each other's scores.

Because of early system implementation difficulties, the phases of the input signal sinusoids were not varied from run-to-run during initial training. Thus, the same input signal was used for all tracking runs, and the subjects had the opportunity of "learning" the input.

Intermediate Training Phase

Each subject participated in a second training phase, and trained under the three conditions which were to be used in the following control phase. One of the conditions ("nominal") was the same as that used during initial training, except that the input signal was randomized from run-to-run, to eliminate possible input signal "learning" effects. The other two conditions, "diffused" and "delayed," also incorporated input signal randomization.

Each subject participated in three training sessions, with each day's session devoted exclusively to one of the three conditions. A balanced design across subjects was chosen to compensate for possible sequential learning

effects. Each session consisted of four three-run blocks, to yield 12 training runs per condition per subject.

Control Phase

Each subject participated in the third (control) phase of the experiment, directed at comparing tracking performance under the three conditions. The tracking task and the three conditions were identical to those used in the previous training phase, and only the order of presentation and number of replications were different. This control phase was completed in a single session of four three-run blocks, with each condition presented once per block. A balanced design was chosen to minimize possible within and between block ordering effects, and yielded four replications per subject per condition.

3.3 Data Processing

The primary data obtained from the tracking runs consisted of tracking scores (errors, stick, and stick rate), describing functions (gain and phase), and normalized remnant.

Tracking scores were calculated for each individual run, and consisted of the variance of the display error, stick, and stick rate. The latter two scores were obtained in units of equivalent plant command (cm and cm/sec), by multiplying the recorded stick signal (in units of lbf) by the stick gain (cm/lbf). Since the stick rate signal was not saved during a tracking run, it was computed from the stick history, using a frequency domain differentiation technique (see Zacharias and Levison (1978) for details).

Describing functions were calculated for each control phase run. Gain and phase values were calculated at each frequency contained in the input disturbance, by dividing the Fourier transform of the stick signal by the transform of the display error. Since the digital implementation introduced a 32 msec phase shift in the measurements, all human operator describing functions were corrected for this shift.

Phase measurements associated with the time delay case were also adjusted by the addition of a 40 msec lag, corresponding to the 40 msec dead-time present in the plant dynamics used in this case. In this way the time delay was associated with the human operator rather than the plant dynamics, so that the delayed condition simulated the effect of any stressor serving to increase the operator's processing time by 40 msec.

Estimates of normalized remnant were obtained by partitioning the equivalent stick spectrum into input-correlated and remnant-related components, and, at each input disturbance frequency, calculating the equivalent remnant power referred to display error rate, and normalizing by the variance of the display error rate. The mathematical details are discussed in Levison et al (1969).

The input-correlated and remnant-related components of the stick spectrum were also used to calculate the correlated-to-remnant power ratio (C/R) at each input disturbance frequency. If a C/R ratio was found to be less than 6 dB (indicating that more than 20% of the response power at that frequency was uncorrelated with the input) the corresponding describing function gain and phase values were excluded from the data base.

4. EXPERIMENTAL RESULTS

4.1 Results of Initial Training Phase

Each subject successfully completed 48 training runs. Because of initial difficulty with the task, subjects occasionally tracked at an input disturbance RMS level of 0.5 cm, rather than the nominal 1.0 cm. To provide a fair base of comparison with scores obtained under nominal conditions, scores obtained with a 0.5 cm RMS input were doubled under the assumption of approximate operator linearity. A total of 24 such run scores were adjusted, out of a total number of 288 runs completed by the six-subject population.

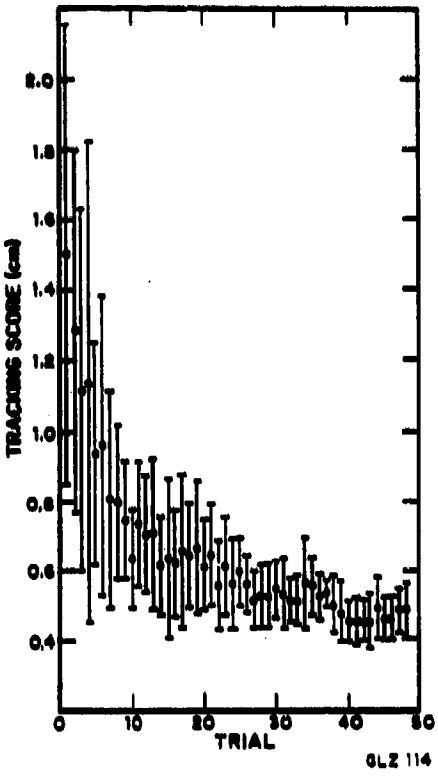
Figure 5 shows tracking score dependence on run number, averaged across the six subjects. Population means and standard deviations are indicated by the dots and bars, respectively. Although 48 runs per subject were conducted, the figure shows a substantial reduction in tracking score by about he 15th trial; by the 30th trial, tracking score was within 20% for the scores obtained at the end of training.

4.2 Results of Intermediate Training Phase

Each subject completed 36 intermediate training runs, 12 on each condition. It was anticipated that training times under all three conditions would be considerably shorter than that seen during initial training. This is confirmed by the tracking score histories shown in Figure 6. The nominal training data show no evidence of a "learning" curve, and the diffused and delayed conditions show a relatively short learning period, on the order of 5 runs.

4.3 Results of Control Phase

Each subject completed 12 control runs, four on each condition. Tracking scores and frequency domain measures were calculated for each tracking run, and averaged across the four replications completed by each subject under each condition. Means were then averaged across subjects, to obtain overall statistical measures of performance with condition. To test for significant differences between conditions, paired differences were formed from corresponding subject means, and subjected to two-tailed t-tests.



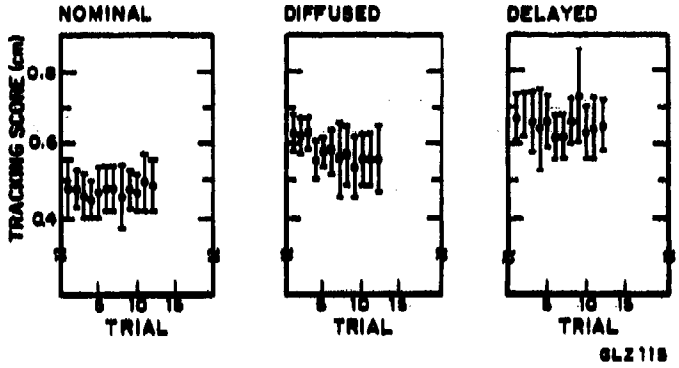


Figure 6: Tracking Score Histories (six subject average)

Figure 5: Tracking Score History During Initial Training (six subject average)

Figure 7 shows score variations due to the differing task conditions. Although the standard deviations are comparable across conditions, the score means show an increase of about 30% over nominal, when subjects track under the diffused or delayed conditions. Similar trends are seen with the stick and stick rate scores. The table shows that tracking score is significantly different across conditions, as is the stick score; the stick rate score, however indicates a significant difference only when nominal tracking is compared with delayed tracking.

Figure 8 summarizes the frequency domain measures associated with the nominal tracking condition; means and standard deviations for the six-subject population are indicated by the error bars at each frequency. The gain data show integrator-like response at low frequencies, and a resonance at approximately 20 rad/sec. The across-subject variance in gain is least in the midband, with larger deviations at both ends of the measurement spectrum. The phase data show a nearly constant lag in the neighborhood of 40 deg, at the low- and mid-frequencies, and then drop off rapidly in the neighborhood of the resonant frequency. Except for the three highest measurement frequencies, the

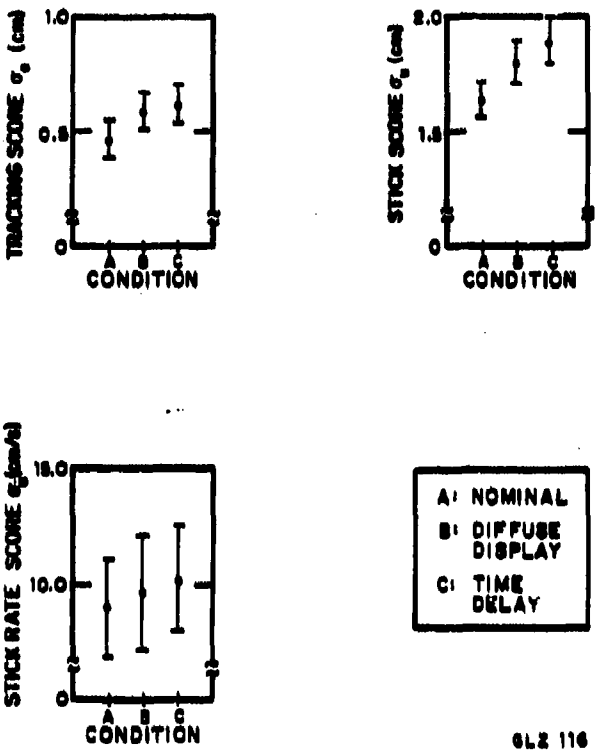


Figure 7: Tracking Scores as a Function of Training Condition (six subject average)

Significance Levels Between Conditions

score \ condition pair	nominal / diffused	nominal / delayed	diffused / delayed
σ_e	.001	.001	.05
σ_u	.001	.001	.01
σ_u^*	N.S.	.001	N.S.

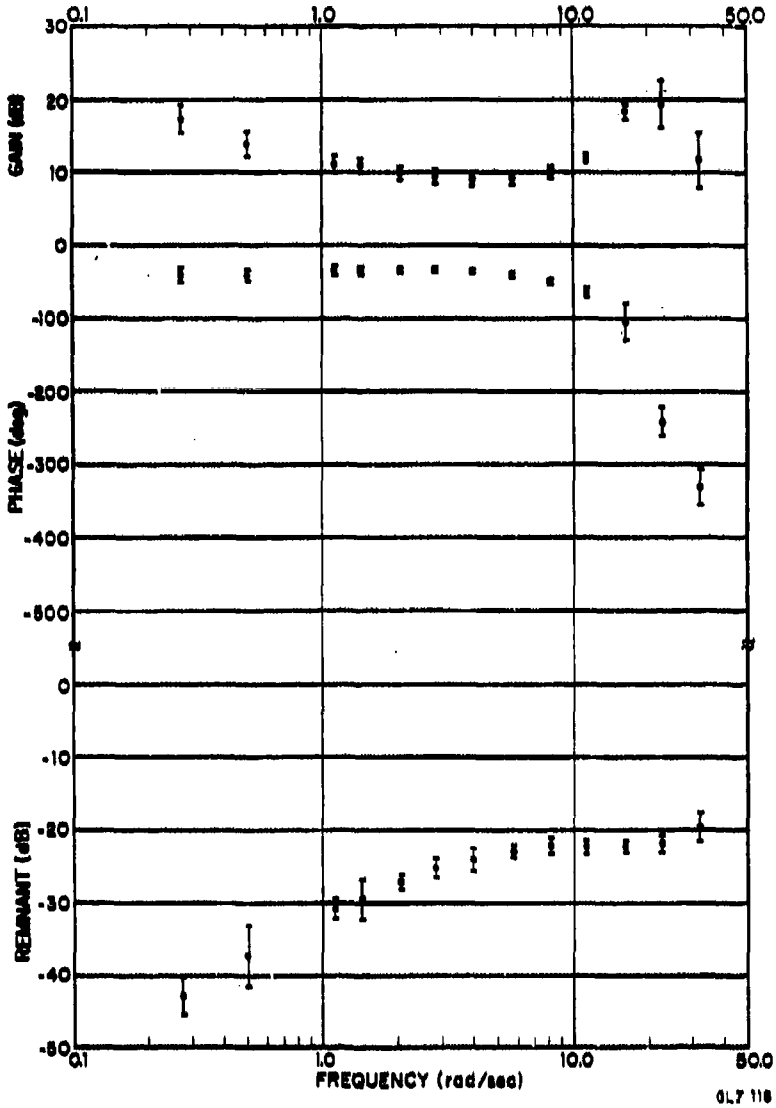


Figure 8: Frequency Response Measures Obtained Under Nominal Condition.

phase data are characterized by very small across-subject variances. The normalized remnant data show a distinctive "washout" pattern over the measurement frequencies, with a break frequency of approximately 4 rad/sec. Except at the lowest measurement frequencies, where there are fewer frequencies over which to do remnant averaging, the data are characterized by very small across-subject variances.

Similar trends with frequency were found for the data associated with the diffused and delayed conditions. Figures 9 and 10 provide a direct pairwise comparison of the mean frequency response measures obtained under the three different conditions, along with the significance levels associated with the differences. (Means having no associated significance symbol were found to be not significantly different, at the 0.05 level.)

Figure 9 provides a comparison of nominal tracking with tracking under diffused display conditions. At the low- and mid-frequencies, operator gain is significantly lower with the diffused display. In the neighborhood of 10 rad/sec this gain difference becomes insignificant; at 16 rad/sec, diffused display conditions are characterized by a higher gain, which appears to drop off rapidly at higher frequencies. These high-frequency gain trends are consistent with a downward shift of the resonant frequency, when going from a nominal display to the diffused display.

The phase data show no significant differences at the low- and mid-frequencies. However, in the region from 8 to 20 rad/sec, the diffused display results in significantly greater phase lags. The lack of a significant phase difference at the highest frequency is consistent with a "rejoining" of the two phase curves.

The remnant data show that although the differences are slight, all of the remnant means are higher with the diffused display, and almost half of the differences are significant. The trends seen are consistent with an overall increase in the remnant level, over the entire spectrum.

Figure 10 provides a comparison of nominal tracking with tracking under delayed conditions. The low- and mid-frequency gains are significantly lower with the time delay present. At 8 rad/sec the difference becomes insignificant, and above that the delayed conditions are characterized by higher than nominal gains. Again, the high-frequency gains trends are consistent with a drop in the resonant frequency, when going from nominal tracking to delayed tracking.

The phase data show significant differences throughout the spectrum, except at the very lowest frequency which has associated with it relatively large variance measurements, and in the neighborhood of 5 to 8 rad/sec. Although not apparent from the figure, the low-frequency lags are less with the time delay present, and thus lead the nominal data in this region. The high-frequency lags are clearly greater with the delay present, and thus the phase relations are

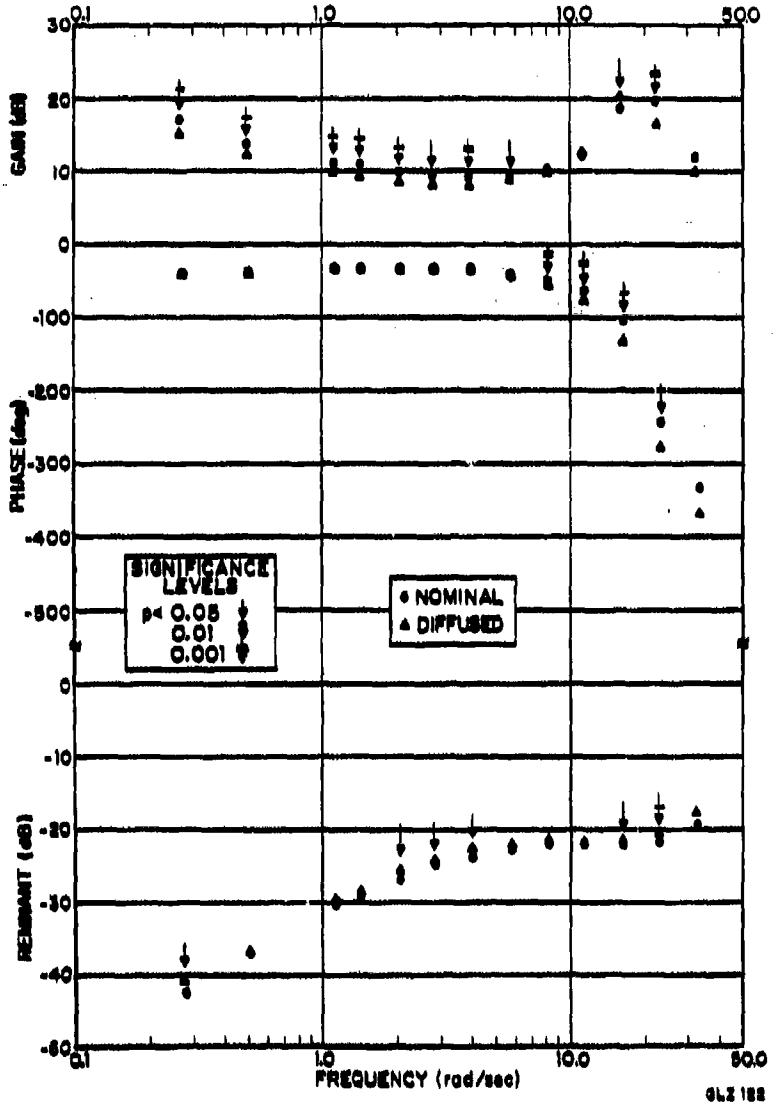


Figure 9: Frequency Response Measures Obtained Under Nominal and Diffused Display Conditions.

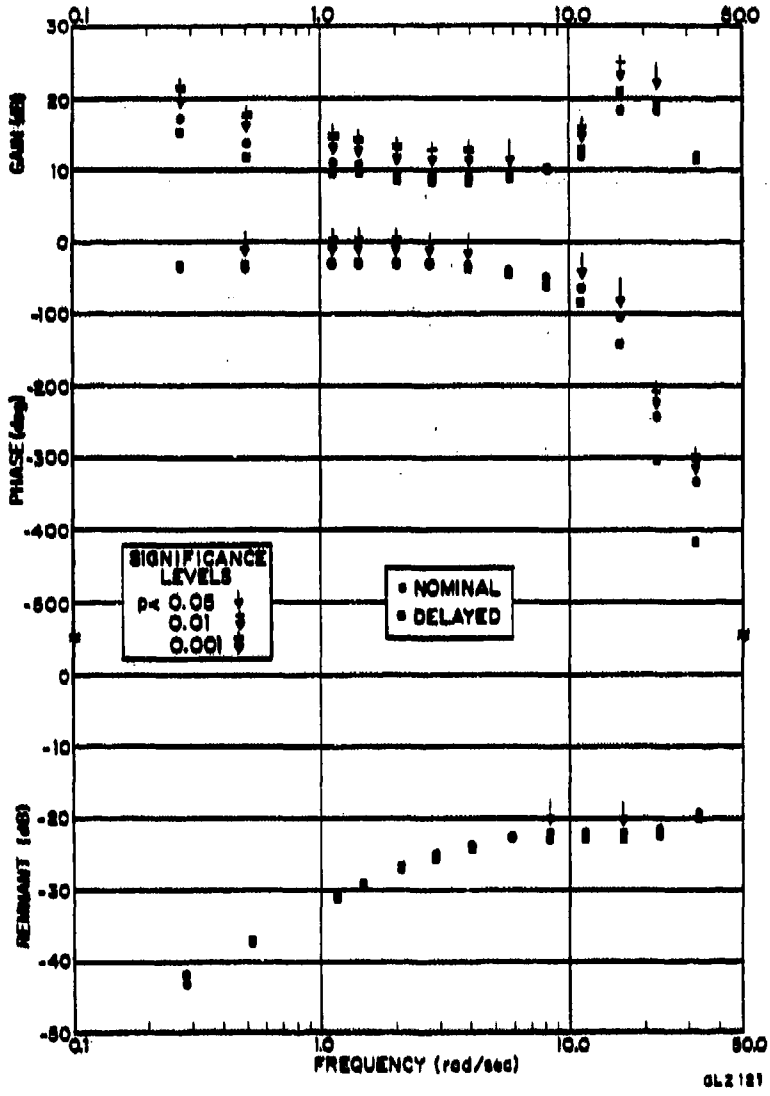


Figure 10: Frequency Response Measures Obtained Under Nominal and Time Delay Conditions.

reversed from that seen at the low frequencies. The 5 to 8 rad/sec neighborhood in which no significant differences are seen is the lead-lag transition region.

The remnant data show no obvious differences at the low- and mid-frequencies; at the high frequencies, the delayed tracking condition is characterized by lower than nominal remnant levels. The fact that only two of the means are (marginally) significantly different suggests that the time delay has little effect on the measured remnant levels.

4.4 Discussion of Experimental Results

The initial training phase showed that the most significant reduction in tracking score occurred within the first 15 trials, and that by the 30th trial, scores were within 20% of their asymptotic levels. Across subject variance showed a similar reduction with run number. This suggests that subjects can learn the task relatively rapidly, and that consistent results can be obtained and maintained with a relatively small subject population, and with a small investment in training.

The intermediate training phase confirmed the asymptotic performance of the subjects on the nominal task, and was characterized by considerably shorter training times on the two stressed tracking tasks. Asymptotic performance on the two stressed conditions was reached in approximately five runs, suggesting that once the nominal task is learned by the subjects, asymptotic performance under off-nominal conditions can be attained relatively quickly.

The control phase showed that tracking error and stick displacement scores were significantly higher under the two stressed conditions, with increments on the order of 30% over nominal. Stick rate score differences were not significantly different, except when comparing nominal with delayed tracking.

Frequency response measures showed the same basic trends with frequency, for all three tracking conditions. A comparison of the measured frequency response given in Figures 9 and 10, with the model predictions of Figures 3 and 4, shows that the model predicts the basic trends with frequency, with a nominal operator parameter set. Since the measured response under the stressed conditions differs only slightly from that made under nominal conditions, the implication is that the model parameters need only be changed slightly from their nominal values, to account for the effects of either stressor.

Four basic features characterize the changes in the frequency response measures when going from the nominal tracking condition to the diffused condition (see Figure 9): a) a lowering of the low- and mid-frequency gain; b) a decrease in the resonant peak frequency; c) an increase in the high-frequency phase lag; and d) an overall increase in remnant level. A review of the model's predicted frequency response, given in Figures 3 and 4, show that these changes can be interpreted as simply an increase in the overall observation noise/signal

ratio. Although variations in other parameters might account for the describing function changes, none of them produce the observed overall increase in remnant level. Thus, if one parameter were to be identified with the changes observed with the diffused display condition, the best candidate would be the overall observation noise/signal ratio. The fact that an increase in this noise level parameter can qualitatively account for the changes seen with diffused display tracking is intuitively satisfying, since one might very well expect a diffuse display to introduce visual "noise" into the tracking loop.

Similar changes are seen when going from the nominal tracking condition to the delayed condition (Figure 10), with the exception that remnant level remains effectively unchanged from its nominal level. For the delayed case, one can thus eliminate from consideration an increase in the operator's observation noise/signal ratio. The model's predicted frequency response (Figures 3 and 4) instead argues for a change in the operator's time delay, the only parameter which accounts for the observed describing function changes and the lack of changes observed in remnant level. The fact that this parameter is effectively identified from the measurements made under the delayed conditions further serves to verify this comparative approach based on model predictions.

Although the discussion has been somewhat qualitative, one can conclude that the tracking task is capable of differentiating between the effects of two different stressors. The approach requires only a brief comparison of the measured frequency response with the human operator model predictions, to isolate the model parameter change most likely responsible for the measured change under the stressed tracking condition. It is anticipated that such an approach would be quite satisfactory for a screening evaluation of tracking stressors; however, it is the intent here to provide a more rigorous justification for this approach, and the next section addresses the problem in a more quantitative manner.

5. MODEL ANALYSIS

A detailed model analysis of the data was conducted using the optimal control model (OCM) of the human operator. The primary objective was to express stressor effects in terms of model parameter changes, so as to provide a quantitative basis for the arguments made above. A secondary objective was to demonstrate the model's utility in reducing the amount of data needed to characterize tracking performance, by condensing a large-dimension data vector to a small-dimension model parameter vector.

5.1 Analysis Procedure

The method for identifying model parameters was similar to that used in earlier study programs and described by Levison et al (1976). Parameter values were sought which would match the various metrics of the model to the

corresponding experimentally derived metrics: performance scores, describing function gain and phase, and remnant level. The matching was accomplished by minimizing a scalar cost function obtained by summing component matching errors, where each component was computed from the sum of the normalized squared differences between model metric and experiment mean; normalization was accomplished by dividing the difference by the associated experimental standard deviation.

The model matching effort was conducted for each of the three data sets obtained under the three conditions tested in the experimental control phase: nominal, diffused, and delayed. No attempt was made to hold particular parameter values constant across conditions; instead, a global search of model parameter values was made for each data set, to arrive at three separate parameter sets, one for each condition.

The optimal control human operator model used in this analysis is described in detail by Levison et al (1976). As in the earlier task design phase, the model was used in a simulation of the tracking task shown in Figure 1. The 32 msec dead-time associated with the digital implementation was simulated by cascading a first-order Pade with the plant dynamics, as was the 40 msec additional dead-time associated with the time delayed condition.

5.2 Analysis Results

Table 1 presents the parameter values obtained from the model matching effort, for the three experimental conditions. The last row of the table gives the resulting normalized matching score and shows that the matching errors range between 0.5 and 0.7 standard deviation.

An increase in time delay is one of the primary effects seen with either of the two stressed conditions: the diffused display results in an 11% increase over the nominal condition delay, and the time delay results in a 23% increase.

The observation noise/signal ratios show that a best fit is obtained when rate information is considered to be a noisier variable than position information, by a consistent 2 dB across conditions. More relevant to the stressor identification problem, however, is the fact that the delay condition is characterized by an observation noise/signal ratio identical to that obtained under nominal conditions. In contrast, the diffused condition results in an overall noise level 1 dB higher.

The parameters related to the "motor" functions of the operator model show little or no change with experimental condition, as might be expected from the nature of the stressors. The pseudo motor noise parameter values shown indicate that the delayed condition is best characterized by a -28 dB level, rather than the nominal -30 dB level. However, the model matching score is relatively insensitive to a change in this parameter, since the 2 dB difference accounts

Parameter	Dimension	Nominal	Diffused	Delayed
Time delay	sec	0.154	0.171	0.189
Observation noise (position)	dB	-25.0	-24.0	-25.0
Observation noise (rate)	dB	-23.0	-22.0	-23.0
Motor time constant	sec	0.080	0.080	0.080
Motor noise	dB	*	*	*
Pseudo motor noise	dB	-30	-30	-28
Matching score	-	0.52	0.50	0.67

* negligible

Table 1: Operator Model Parameters for Three Experimental Conditions

for less than 2% of the matching error. Thus, the three conditions can be reasonably matched by an identical -30 dB pseudo motor noise/signal ratio.

Although the matching scores of Table 1 give some indication of "goodness-of-fit," a more graphic illustration is provided by a direct comparison of model predictions with the experimental data. Figure 11 repeats the performance scores of Figure 7; the superimposed open circles indicate the model predictions obtained from the matching procedure just described. All of the trends with condition are closely followed, with small matching errors.

Figure 12 repeats the nominal condition frequency domain data presented earlier in Figure 8; the smooth curves indicate the model predictions obtained from the matching effort. All of the frequency trends are closely followed in this nominal case; similarly close matches were obtained with the diffused and delayed data sets.

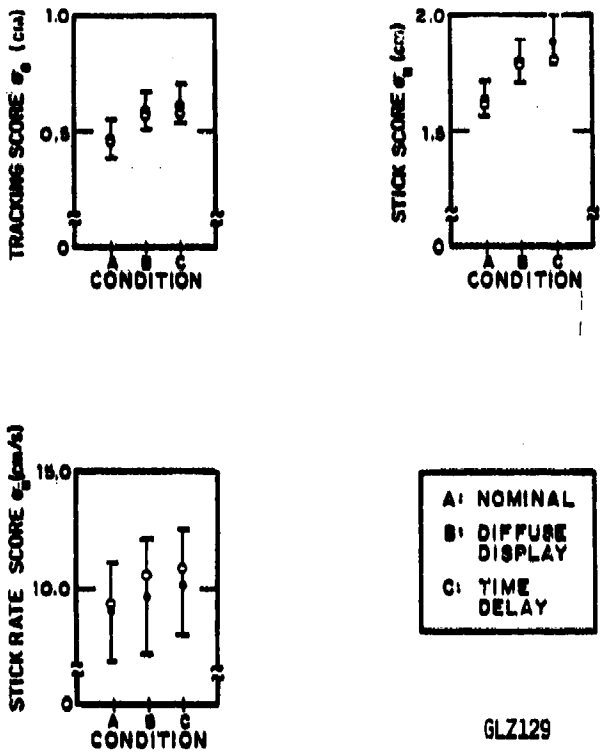


Figure 11. Comparison of Model and Experimental Performance Scores (model - open circles)

5.3 Discussion of Analysis Results

These results demonstrate that a considerable amount of data compression can be achieved by suitably adjusting the model parameters to fit the data. Each data set consisted of 84 measurements (means and standard deviations of the scores and frequency domain measurements); the matching procedure effectively reduced these to 5 specified model parameters, and managed to fit the data within approximately 0.5 standard deviations, on the average.

Of direct relevance to the stressor identification effort is the task's demonstrated ability to provide the experimenter with a means of differentiating between stressors.

Table 1 shows that the diffused display had two effects: a) a 1.0 dB increase in the operator's overall noise/signal ratio; and b) a 17 msec increase

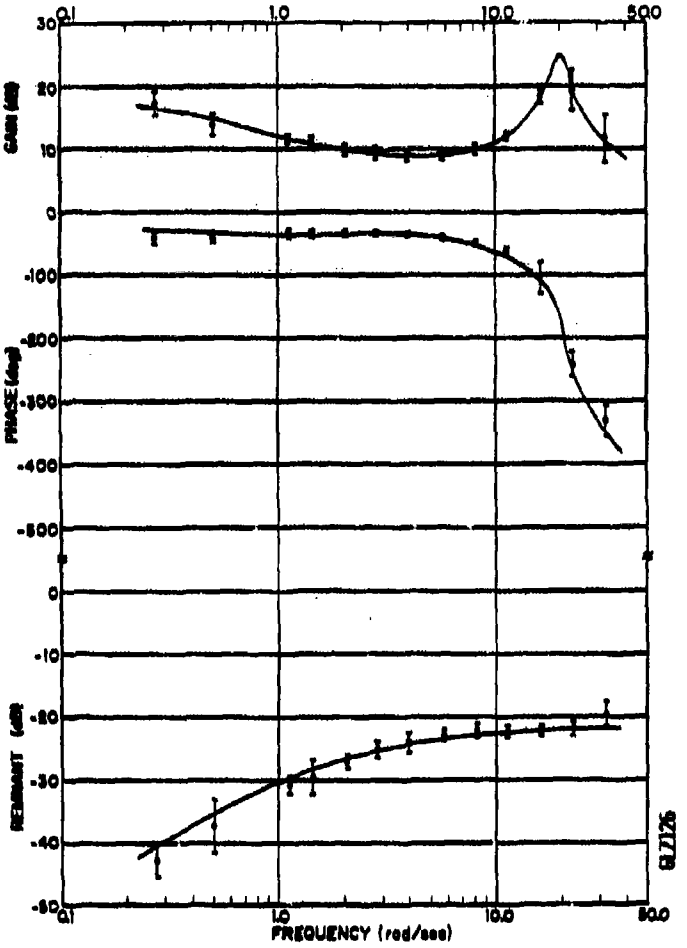


Figure 12: Comparison of Model and Experimental Frequency Measures - Nominal Condition.

in his time delay. All other parameters were unchanged from the parameter set associated with nominal tracking conditions. The implication is that the diffused display not only added "visual" noise to the tracking loop, but also was responsible for increasing the operator's time delay. The first effect is consistent with the characteristics of the diffuser, and serves to validate the utility of both the tracking task and the model analysis in identifying stressor characteristics. The second effect was not expected, since it was not anticipated that additional processing time would be needed to process a "noisy" display signal. The model analysis results, however, indicate that additional processing time was indeed associated with the stressor, and that it is inappropriate to assume a one-for-one correspondence between an apparent stressor characteristic (e.g., "noiseness") and a change in a single human operator model parameter (e.g., observation noise/signal ratio).

Table 1 also shows that the time delay had two effects on the operator: a) a 35 millisec increase in the operator's time delay; and b) a 2 dB increase in his pseudo motor noise/signal ratio. All other parameters were unchanged from the parameter set associated with nominal tracking conditions. As noted earlier, the 2 dB noise increase is of questionable significance, since it only contributes to 2% of the matching error. The 35 msec increase in the operator's time delay, however, is significant, and corresponds closely to the 40 msec delay added to the loop, a delay which effectively simulated the effect of any stressor serving to increase only the operator's time delay. The fact that the model analysis accounted for 35 of the 40 msec, with no call for a readjustment of other operator parameters, again validates the task's ability in identifying stressor effects.

Because the model fitting procedure is highly non-linear, no variance estimates were available for the derived model parameters. Thus, parameter value confidence intervals could not be generated, and statistical tests of parameter differences between conditions could not be performed. The conclusions made above are thus subject to a more rigorous statistical verification, which must wait the development of an algorithm for estimating model parameter variances.

6. SUMMARY

The principal results of this study may be summarized as follows:

1. The use of an unstable controlled element provided the desired task sensitivity to environmental stress, and the incorporation of a sum-of-sines disturbance signal allowed for the calculation of important operator frequency domain measures required for the interpretation of stressor effects.

2. Subjects learned the basic tracking task relatively rapidly, and in 15 runs were relatively close to their final asymptotic scores. After learning the basic task, subjects required on the order of 5 additional runs to reach approximate asymptotic performance under the stressed conditions.

3. The task demonstrated sensitivity to both stressors used in this study: a display diffuser which degraded the target image, and a 40 msec dead-time which delayed the operator's control. Tracking scores under both stressed conditions were significantly higher than that obtained under nominal conditions.

4. The describing function and remnant measures obtained from the tracking data showed significant differences between experimental conditions, and verified the trends predicted by the pre-experimental simulations conducted with the human operator optimal control model. A direct comparison between the data and the model simulations showed that the diffuser acted to increase the operator's observation noise/signal ratio; in contrast, the dead-time acted to increase the operator's time delay. The task thus demonstrated a differential sensitivity to the stressors, and the chosen score and frequency domain metrics allowed for an identification of specific stressor qualities.

5. A quantitative model analysis validated the qualitative conclusions regarding stress effects on tracking behavior. A model fit to the diffused display data showed a 1.0 dB higher than nominal observation noise/signal ratio, and a 17 msec longer than nominal operator time delay, results which are consistent with a "noisy" visual image generated by the diffuser. The dead-time tracking data was best fit with a 35 msec longer than nominal operator time delay, again an effect consistent with the quality of the stressor.

REFERENCES

1. Jex, H. R., J. D. McDonnell, and A. V. Phatak, "Critical Tracking Task for Man-Machine Research Related to the Operator's Delay Time," NASA CR-616, 1966.

2. Kleinman, D. L., S. Baron, and W. H. Levison, "An Optimal Control Model of Human Response - Part I: Theory and Validation," Automatica, 6: 357-369, 1970.

3. Levison, W. H., S. Baron, and A. M. Junker, "Modeling the Effects of Environmental Factors on Human Control and Information Processing," AMRL-TR-76-74, Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, August 1976.

4. Levison, W. H., S. Baron, and D. L. Kleinman, "A Model for Human Controller Remnant," IEEE Transactions on Man-Machine Systems, MMS-10: 101-108, December 1969.

5. Replogle, C. R., et al., "Human Operator Performance in Hypoxic Stress," Aerospace Medical Research Laboratory, Wright-Patterson Air Force Base, Ohio, November 1970.

6. Zacharias, G. L., and W. H. Levison, "A Performance Analyzer for Identifying Changes in Human Operator Tracking Strategies," BBN Report No. 3910, Bolt Beranek and Newman Inc., Cambridge, Mass, August 1978.

ACKNOWLEDGEMENTS

The work reported here was performed under Aeronautical Systems Division/PPMSS contract number F33615-77C-0522. Technical Monitor for the Air Force was Lt. Col. Robert D. O'Donnell of the Aerospace Medical Research Laboratory (HEA).

OPTIMALIZATION AND EVALUATION OF LINEAR MOTION FILTERS

by
Ruud J.A.W. Hosman
Hans C. van der Vaart
Gerrit A.J. van de Moesdijk
Delft University of Technology
Department of Aerospace Engineering
Kluyverweg 1
Delft - The Netherlands

SUMMARY

A time-domain optimization computer routine was designed to find optimum motion filter settings for the pitch and heave motion in a three degrees of freedom (pitch, heave and roll) flight simulator. Closed loop aircraft dynamics of an altitude holding task in a light aircraft in turbulence, including a pilot model, were used in the computer runs. Cost functions were based on the difference between vestibular outputs in the moving base simulator and in actual flight, extra penalties being included for the occurrence of false cues and the exceedance of motion system actuator limitations.

Some promising filter settings, together with a few limit-cases were evaluated in a piloted simulator experiment. Both pilot opinion and pilot performance parameters appeared to be significantly better for the particular limit case that gave a one to one simulation of pitching rotations, at the cost of completely neglecting the specific forces associated with the aircraft's longitudinal accelerations.

1. INTRODUCTION

In the field of flight simulation, a trend towards an ever more realistic imitation of actual flight conditions can be observed. For a high quality of simulation all sensory cues that give information to a pilot controlling an aircraft have to be imitated as true to nature as possible, whether they be aural, visual, tactile or vestibular. Especially the visual simulation of the flight environment can be very sophisticated today using multiscreen, coloured and textured computer generated images of the outside world. The visual illusion of taking off from one airfield, of climbing and cruising flight followed by approach and landing at a different airfield can almost be perfect in a present day simulator. However, if vestibular cues (rotational accelerations and specific forces) are considered, such a perfect imitation of actual flight can never be attained. Although certain angular displacements and rotational accelerations, especially those occurring in every day airline flight can still quite realistically be simulated, a true simulation of specific forces due to linear acceleration as occurring in real flight can never be obtained due to the limited travel of simulator actuators. Moreover very often still more limitations are set if a simulator motion system has a limited number of degrees of freedom. All these limitations set by the physical constraints of motion systems, call for the use of motion (or washout) filters and the resulting motion is necessarily always some form of compromise, see for instance Refs. 1, 2 and 3.
Apart from the constraints set by the motion system, the choice of the gains and

parameters of the motion filters is also influenced by the particular type of aircraft to be simulated and the piloting task to be accomplished. For instance a take-off roll, a tracking task in atmospheric turbulence or manoeuvres in still air all call for different motion filtering.
In Ref. 4, for instance, the distinction between "manoeuvre motion" such as occurring when performing a banked turn, a side step manoeuvre or a landing flare in still air on the one hand, and "disturbance motion" due to for instance atmospheric turbulence on the other, has been stressed.

At the Department of Aerospace Engineering of the Delft University of Technology research is done in the field of pilot behaviour, handling qualities and flight simulation. Part of this research is directed to the perception process of a pilot controlling an aircraft. Experiments on thresholds of motion perception (Ref. 5) proved that the Department's three degrees of freedom flight simulator with its unique high fidelity motion system, see Ref. 6, is a valuable tool for research in this field.
Studies of human sense organs have shown that there are narrow relations between some of the senses, especially between vision on the one hand and vestibular sensations on the other. Such relations and interactions must have an important influence on pilot's perception process, see Ref. 7. Moreover, it is known that some of the outputs of the different sense organs result in a redundancy in information. This is a promising aspect for flight simulation engineers as it should be possible to obtain important savings by limiting one or more motion components to be simulated without seriously imparing the total perception process and, hopefully, the quality of the simulation.
Unfortunately, the entire perception process is not fully understood yet and it is difficult to decide which particular motion component might be limited and in what manner. The only practical solution at present is to generate the aircraft's motion, as sensed by the vestibular organs, as accurately as possible.

The present paper considers one limited aspect of a particular case of motion simulation. It deals with the symmetric motions in atmospheric turbulence of a low wing loading light aeroplane.
A number of different motion filter configurations were optimized using a time-domain computer simulation of an altitude holding task in atmospheric turbulence. The constraints were those set by the limitations of the Aerospace Department's moving base flight simulator.
Although a computer simulation was also done for the case of manoeuvre motion in still air, a piloted simulator experiment was not done for that case. The experiments described in this paper only concern the altitude holding task in which the manoeuvre and disturbance motion, just mentioned, are both included. After the theoretical optimization process, a number of most promising filter configurations and settings were evaluated in a piloted simulator experiment.

2. MOTION FILTER OPTIMIZATION

2.1. Design of the computer simulation

The aim of the computer simulation was to obtain motion filter settings that would result in a best possible simulation of the aircraft symmetrical motions as sensed by the pilot's vestibular system, given the physical limitations of the simulator's motion system (limited numer of degrees of freedom and maximum excursions, velocities and accelerations).

The inclusion of the dynamics of the vestibular system in the computer simulation

means that only errors in angular accelerations and specific forces as sensed by the pilot's vestibular system are to be minimized. Due to the dynamics of the vestibular system certain errors in the simulated motion will hardly be perceived. It would be irrational to try to obtain a quality of motion any better than a human pilot would be able to sense. Therefore it was decided to optimize the motion filters for minimum difference between the vestibular outputs of the pilot in the aircraft and in the simulator.

It should be remarked that the vestibular tresholds of angular acceleration and those of specific force perception have been left out of consideration. These thresholds have been shown to be in the order of $0.05^{o}/sec^{2}$ for rotations and 0,01 g for specific forces, see Ref. 5. These values are so low that in practice it is quite impossible to design motion (or washout) filters such that any false cues would remain below thresholds.

A schematic representation of the computer simulation is given in Fig. 1. The simulation concerns the motion filters for symmetric aircraft motions in a simulator having only pitch and heave motion for symmetric motion simulation.

As can be seen from Fig. 1, the optimization of the motion filter characteristics by minimizing the error E(t) depends on the aircraft motions and thus on the dynamic characteristics of the particular aircraft type concerned, on the external disturbances and on the task to be performed.
In the case of the present study the combined effects of manoeuvre and disturbance motion were included.
Also from Fig. 1, it can be seen that the pilot model lacks a feedback of the vestibular output. Vestibular information definitely helps a pilot in performing his control task, as will become evident in this paper. The pilot-aircraft system in this computer simulation was only used to produce reasonable aircraft state time-histories.

The aircraft to be simulated was the De Havilland DHC-2 Beaver, the Aerospace Department's laboratory aircraft. Its dynamic characteristics are accurately known, see Ref. 8, and the subjects participating in the piloted simulator experiment were familiar with this aircraft type. Due to the low wing loading this aircraft is rather sensitive to turbulence, making it a typical "difficult" one to simulate realisticly.

The piloting task in the computer simulation as well as in the piloted simulator experiment consisted of keeping a reference altitude in a gusty atmospheric environment.

2.2. Motion perception

The perception of motion is usually modelled as exclusively sensed by the vestibular system, consisting of semicircular canals, sensitive to angular accelerations, and otoliths, sensitive to specific forces. For an extensive description the reader is referred to Ref. 5.

The dynamics of the semicircular canals and the neural preprocessing within the organ can be modelled by the following transfer function relating the vestibular output R to the angular acceleration $\ddot{\theta}$:

$$\frac{R(s)}{\ddot{\theta}(s)} = \frac{1 + 0.11\ s}{(1 + 5.9\ s)\ (1 + 0.005\ s)} \qquad (2.1.)$$

As concerns the otolith, recent experiments in the field of neuro-physiology on the dynamic behaviour of the sensory cells in the otoliths yielded the following transfer function relating the vestibular output S to the input specific force A:

$$\frac{S(s)}{A(s)} = \frac{1 + 10\ s}{(1 + 5\ s)\ (1 + 0.01\ s)} \qquad (2.2.)$$

Eqs. (2.1.) and (2.2.) were used to compute vestibular output in the computer simulation. The small time constants in the denominators of eqs. (2.1.) and (2.2.) have been neglected.

2.3. Motion filters

For the case of symmetric flight considered here the aircraft's equations of motion are:

$$A_X - g \sin\theta = \dot{u} - \dot{\theta}w$$

$$A_Z + g \cos\theta = \dot{w} - \dot{\theta}u$$

$$M = I\ddot{\theta}$$

The vestibular system is sensitive to angular accelerations and specific forces. In the aircraft the inputs to the vestibular system, neglecting the small term $\dot{\theta}w$, will be:

$$A_X = g \sin\theta + \dot{u}$$

$$A_Z = -g \cos\theta + \dot{w} - \dot{\theta}u \qquad (2.3.)$$

$$\ddot{\theta} = \frac{M}{I}$$

The purpose of the present optimisation procedure was to simulate the vestibular outputs S_X, S_Z and R_Y as true to the actual values as possible under the constraint that angular and linear simulator accelerations, velocities and displacements remain within the limitations of the motion system. During routine airline flight the angular displacements (especially of the angle of pitch θ) are roughly within the angular limitations of motion systems. This is evidently not the case with the

linear displacements. It turns out that, although the variations of specific forces during flight are well within the capabilities of most motion systems, a primary limitation of all motion systems is the one due to limited linear displacements. Hence the possible duration of constant accelerations in a simulator is rather short.
If it is assumed that the term g cos θ in the vertical specific force A_Z (see eq. (2.3.)), is taken into account by the gravity component in the simulator g cos θ_s then the normal acceleration:

$$a_n = -A_Z - g \cos \theta$$

$$= -\dot{w} + \dot{\theta} u \qquad (2.4.)$$

has to be generated by the simulator's heave motion. According to Ref. 1, a high pass filter will be used having the transfer function:

$$H_{H_n}(s) = \frac{(\tau_{H_n} s)^2}{(1 + \tau_{H_n} s)^2} \qquad (2.5.)$$

In this way long duration of constant normal accelerations, which would lead to large vertical displacements, are prevented.

In the present case of a three degrees of freedom (pitch, heave and roll) motion system, only simulator pitch θ_s is available for generating both pitching acceleration and longitudinal specific force cues. Two basic filter configurations to generate these cues have been considered.

Filter configuration A:high pass filtering of θ, low pass filtering of A_X
This filter configuration, which is the one most commonly used, see Ref. 1, is depicted in the diagram of Fig. 2. The simulator pitch angle θ_s is generated by the sum of two filters. The high pass filter with input θ controls the simulator pitch angle in the high frequency range:

$$H_{H_\theta}(s) = \frac{(\tau_{H_\theta} s)^2}{(1 + \tau_{H_\theta} s)^2} \qquad (2.6.)$$

The low pass filter with input $\frac{A_X}{g}$ is generating the low frequency pitch movement of the simulator:

$$H_L(s) = \frac{1}{(1 + \tau_L s)^2} \qquad (2.7.)$$

Since the values of θ remain small the input to the low pass filter is taken to be:

$$\frac{A_x}{g} = \sin\theta + \frac{\dot{u}}{g} \approx \theta + \frac{\dot{u}}{g}$$

In this case, see Fig. 2, the simulator pitch angle is:

$$\theta_s(s) = H_{H\theta}(s)\ \theta(s) + H_L(s)\left\{\theta(s) + \frac{\dot{u}}{g}(s)\right\}$$

$$= \left\{H_{H\theta}(s) + H_L(s)\right\}\theta(s) + H_L(s)\ \frac{\dot{u}}{g}(s) \qquad (2.8.)$$

In the optimization procedure the high pass and low pass filters can be adjusted such that the errors in the vestibular responses R_{y_s} and S_{x_s} are minimized.

Filter configuration B: no filtering of θ, low pass filtering of $\dot{u}$
If in the case of filter configuration A the transfer function $\{H_{H\theta}(s) + H_L(s)\}$ is approximately equal to one over the frequency range of interest, it is preferable to use filter configuration B, see Fig. 3:

$$\theta_s(s) = \theta(s) + H_L(s)\ \frac{\dot{u}}{g}(s) \qquad (2.9.)$$

The filter configurations according to eqs. (2.8.) and (2.9.) have been used in the optimization procedure discussed in the next paragraph.

2.4. Optimization of the motion filters

The time constants of the different motion filters were optimized by minimizing the difference between the output of the vestibular system in actual flight and in the simulator, as determined during a computer simulation run. The block diagram of Fig. 1 gives an impression of this procedure.
Particulars of the aircraft are given in Table 1. The aircraft was perturbed by vertical and horizontal turbulence with spectra according to Dryden, see Ref. 9. The turbulence velocities had a standard deviation of 1 m/sec, this kind of turbulence can be termed as light, see Table 2.
The pilot model controlling the aircraft was taken from Ref. 10.

$$H_p(s) = K_p\ \frac{1 + T_1 s}{1 + T_I s}\ e^{-Ts}$$

The model parameters were adjusted such that good response characteristics to

gust disturbance and to commanded altitude profiles were obtained. The cross-over frequency and phase margin of inner and outer loop and the model parameters are given in Table 3.
The cost function to be minimized by varying the motion filter time constants, was based on:

$$J = \int_{t_o}^{t_1} \left\{ W_1 \frac{(S_{X_a} - S_{X_s})^2}{\sigma^2_{S_{X_a}}} + W_2 \frac{(S_{Z_a} - S_{Z_s})^2}{\sigma^2_{S_{Z_a}}} + W_3 \frac{(R_{Y_a} - R_{Y_s})^2}{\sigma^2_{R_{Y_a}}} \right\} dt \quad (2.10.)$$

S_X and S_Z are the responses of the otoliths due to the input specific forces A_X and A_Z; R_Y is the response of the semicircular canals due to the input angular acceleration $\ddot{\theta}$. The subscripts a and s denote values in aircraft and simulator respectively. The computer runs covered 115 secs from t_o to t_1.

Due to the washout characteristics of the motion filters it is possible that one or more of the components of the vestibular outputs in the simulator have a sign opposite to that in the aircraft. Such false cues have a very disturbing effect on a pilot controlling the simulator. In order to limit the occurrence of false cues the weighting coefficients W in eq. (2.10) were made to depend on the sign of the vestibular outputs:

$W_1 = 1$ if $S_{X_a} \cdot S_{X_s} \geq 0$ $\quad$ $W_1 = 4$ if $S_{X_a} \cdot S_{X_s} > 0$

$W_2 = 1$ if $S_{Z_a} \cdot S_{Z_s} \geq 0$ $\quad$ $W_2 = 4$ if $S_{Z_a} \cdot S_{Z_s} > 0$

$W_3 = 1$ if $R_{Y_a} \cdot R_{Y_s} \geq 0$ $\quad$ $W_3 = 4$ if $R_{Y_a} \cdot R_{Y_s} > 0$

Finally the limitations of the motion system had to be included in the cost function. This was achieved by adding a penalty to the cost function if any of the servo actuators of the motion system were calculated to reach their limits (displacement, speed, acceleration). The value of this penalty was rather arbitrarily chosen but resulted in a considerable rise of the cost function whenever the motion system was exceeding its maximum excursions. It turned out that for the altitude holding task for this particular aeroplane and the intensities set for the atmospheric turbulence is was not necessary to include penalties for simulator actuator velocity or acceleration, as they remained within the motion system's capabilities. See Table 4.
Due to the nonlinear components in the cost function (weighting coefficients penalizing false cues and the penalty for exceeding maximum servoactuator travel) the optimization process was not fit for automatic linear system theory optimization routines. Therefore the error signals were computed as a function of time in a computer simulation using discrete steps of 0.05 secs.
The optimization process was carried out by repeating the computer run a number of times and by making stepwise changes in the filter time constants for every new run.
The dynamic characteristics of the simulator motion system are, during actual simulation, compensated for by a compensating network. See Ref. 6.
These characteristics and their compensation could therefore be deleted in the computer program.

The cost function was obtained by:

$$J = J_1 + J_2 + J_3 + J_4$$

where J_1, J_2 and J_3 correspond with the terms in eq. (2.10.) representing the weighting of the errors in the vestibular output.

$$J_1 = \frac{0.05}{\sigma^2_{s_{X_a}}} \sum_{i=1}^{N} W_1 \, (s_{X_{a_i}} - s_{X_{s_i}})^2 \qquad (2.11.)$$

$$J_2 = \frac{0.05}{\sigma^2_{s_{Z_a}}} \sum_{i=1}^{N} W_2 \, (s_{Z_{a_i}} - s_{Z_{s_i}})^2 \qquad (2.12.)$$

$$J_3 = \frac{0.05}{\sigma^2_{R_{Y_a}}} \sum_{i=1}^{N} W_3 \, (R_{Y_{a_i}} - R_{Y_{s_i}})^2 \qquad (2.13.)$$

Finally J_4 is the component in the cost function penalizing the exceedence of front and aft actuator limitations:

$$J_4 = 0.05 \sum_{i=1}^{N} P_{1_i} + 0.05 \sum_{i=1}^{N} P_{2_i} \qquad (2.14.)$$

P_{1_i} and P_{2_i} are weighting coefficients for the fore and aft moving actuators respectively. They depend on the magnitude of the computed actuator excursions $h_{1_{s_i}}$ and $h_{2_{s_i}}$ in the following manner:

$P_{1_i} = 0$ if $|h_{1_{s_i}}| < .30$ m $P_{1_i} = 10$ if $|h_{1_{s_i}}| \geq .30$ m

$P_{2_i} = 0$ if $|h_{2_{s_i}}| < .30$ m $P_{2_i} = 10$ if $|h_{2_{s_i}}| \geq .30$ m

2.5. Results of the computer simulation

Before going into the results obtained, it should be borne in mind that it is difficult to interpret the results directly in terms of aircraft dynamic behaviour. The calculated cost functions discussed below are a result of the characteristics of the closed loop formed by the aircraft plus human pilot model, see Fig. 1. Effects of aircraft response to horizontal and vertical turbulence, human pilot model response to altitude deviations and aircraft response to

elevator control inputs are all included in the computer simulation.

After a small number of simulation runs it became evident that the motion filters for the simulator pitch control could be considered quite independently from the filter controlling the normal acceleration. Therefore the results are treated separately, first those for the two basic filter configurations for the pitch control followed by the results for the normal acceleration filter.

For filter configuration B (no filtering of θ, low pass filtering of $\dot{u}$), the components J_1 and J_3 (see eqs. (2.11.) and (2.13.) of the cost function) are plotted as a function of the low pass filter time constant τ_L in Fig. 4. The magnitude of J_1, which is an indication of the error in the specific force simulation, is seen to increase with τ_L. The error in perceived angular acceleration as given by J_3, is only caused by simulator rotations due to the simulation of $\dot{u}$, see Fig. 3.

Increasing τ_L or decreasing the low pass filter break frequency, causes a deterioration of the specific force simulation but improves the angular acceleration simulation, leading to a minimum of the sum of J_1 and J_3 at around $\tau_L = 0.27$ secs. Of course this minimum, denoted by F1, depends on the relative weighting of the errors in specific force and angular acceleration response.
As there was no a priori knowledge of the manner in which a human pilot weights these two different vestibular sensations the weighting coefficients W_1, W_2 and W_3 were set at equal values, see par. 2.4.

Increasing τ_L to infinity, which actually means that the longitudinal acceleration is neglected and the aircraft's pitching motion is simulated one to one, will drive J_1 to a final value and J_3 to zero. This filter setting, denoted by F5, was included in the piloted simulator study. It should be remarked that a filter setting with characteristics similar to F5 can be obtained with configuration A if $\tau_{H\theta} = \infty$ and $\tau_L = \infty$.

For the filter configuration A a simular influence of the low pass filter time constant can be observed in Fig. 5. The value of the cost function J_1 for the specific force simulation is, for all settings of the high pass filter time constant $\tau_{H\theta}$, an order of magnitude lower than the function J_3 for the angular acceleration and only the sum of J_1 and J_3 has been plotted for a number of values of $\tau_{H\theta}$ as a function of τ_L.

In earlier similar computer simulation experiments the optimization procedure of the type A filter was stopped after finding a local, weak minimum in the same region as the minimum denoted by F2 and F3 ($\tau_L = 1$, $\tau_{H\theta} = 1$ and $\tau_L = 0{,}4$, $\tau_{H\theta} = 0.5$ secs respectively). It was, at that time, not considered worthwhile to exploit values of $\tau_{H\theta}$ any lower than 0.4 secs. It is plausible indeed, bearing in mind that the aircraft's fast natural symmetrical mode has a frequency around 4 rads/sec for this type of aircraft, that a strong deterioration of the simulator's rotational accelerations could be expected by choosing the high pass filter break frequency any higher than 4 rads/sec ($\tau_{H\theta}$ lower than 0.25 secs).

A surprising decrease in the cost function $J_1 + J_3$ is obtained for the limit case where $\tau_{H\theta} = 0$ ($H_{H\theta}(s) = 0$). This means that the simulator pitch angle is controlled only by the low pass filtering of the aircraft's specific force A_X. The minimum value of the cost function, denoted by F4, is about the same as found for the filter type B (minimum F1).

In order to give an idea of the influence of the different filters on the response of the vestibular system, the responses to a commanded stepwise change of the pitch angle θ are plotted in Fig. 6. From the time-histories of Fig. 6 concerning manoeuvre motion only it is apparent that, due to the fact that the angle of pitch θ and the longitudinal acceleration $\dot{u}$ are roughly of opposite sign, no satisfactory simulation of both θ and A_X can be obtained with the present moving base configuration.
The large differencies between the responses of the selected filters are evident. In Fig. 7 Filter 1 gives good initial responses of θ_s and Ry_s, but the specific force Ax_s and the otolith response Sx_s are much too high. The filters 2 and 3, both configuration A filters, tend to improve the specific force response at the cost of the simulator pitch rotation and corresponding semicircular canals response. The plots of filter 4 appear to be the worst of all although Ry_s is exactly in phase with Ry_a. This is important bearing in mind the weighting coefficients of the cost function. For filter 5, θ_s equals θ, the specific force Ax_s and the otolith output Sx_s, however, are extremely high.

Because it is very hard to choose the most promising filter it was decided to use all these filters in the experiment to evaluate their influence on pilot's performance and opinion. As a reference an additional configuration with fixed simulator pitch angle, termed filter 6, was included in the simulator experiment.

None of the simulator pitch control filter settings were influenced by the penalty set for exceeding actuator limits. These penalties, however, evidently had an influence on the cost functions of the normal acceleration simulation, as can be seen in Fig. 8. The cost function decreases with increasing time constant τ_{H_n}. A sharp rise occurs in $J_2 + J_4$ when the increasing vertical displacements exceed the motion system limitations. Based on the minimum value of $J_2 + J_4$ the time constant τ_{H_n} was chosen at 0.29 secs.

3. EVALUATION OF THE MOTION FILTERS

3.1. The piloted simulator experiment

The six particular settings of the motion filter configurations A and B as described in the preceding chapter were evaluated in a piloted simulator experiment. The motion filter settings are summarized in Fig. 9. Also included is the reference case F6 where there is only simulator heave and no simulator pitch.

The linearized symmetric aircraft motions due to turbulence and control deflections were similar to those of the computer simulation experiment. In order to make the piloting task a realistic one, the asymmetric linearised motions due to turbulence and control deflections were also simulated. The motion filters generating the simulator roll motion were of the standard type used for this particular simulator, see Fig. 10.

Standard flight instruments and controls were available for the piloting task of holding the simulated aircraft at constant altitude.

Due to the fact that the three participating subjects were all type rated Beaver pilots and were, moreover, familiar with the particular simulator, only short training periods were necessary.

A randomised factorial design was used for the experiment. Every subject

performed 5 series of 6 runs, each series having a different, random sequence. The total number of measurement runs was 3 x 5 x 6 = 90, the runs had a duration of 240 secs. The random atmospheric turbulence disturbing the aircraft consisted of the same realisation for each run. During each run seven aircraft variables were recorded: pitch angle θ, rate of pitch $\dot{\theta}$, deviations h from reference altitude, elevator deflection δ_e, yaw angle ψ, roll angle φ and aileron deflection δ_a.

Subjects were asked to give a Cooper-Harper pilot rating for the pitch control of the simulated aircraft after each run. They were also required to give their comments in writing on this aspect of the task.

3.2. Results

The results of the piloted simulator experiment have been evaluated by considering pilot performance parameters, pilot ratings and comments.

Pilot performance parameters

An impression of the subject's performance can be gained by considering the standard deviation of θ, $\dot{\theta}$ and h while the standard deviation of δ_e gives an indication of the pilot effort needed to attain a certain level of performance.

Table 5 and Fig. 11 show significant differences due to the influence of the motion filters. Filter no. 5 (simulation of θ directly, neglecting the simulation of $\dot{u}$) shows the lowest value of σ_θ, $\sigma_{\dot{\theta}}$ and σ_{δ_e}. The results of the analyses of variance are given in table 6.

Further it is remarkable that filter 6, with fixed simulator pitch angle shows a performance as good as filters 1, 2, 3 and 4. Altitude variations do not show any significant changes due to changing filter characteristics.

From the analysis of variance it appears that there are significant differences in the replications. Analysis of the performance expressed by the symmetric variables shows this not to be due to a learning effect. The standard deviation of the symmetric variables of the third replication appeared to be higher than others. This coincides with the unforseen occurrence of sticktion in the simulator altimeter observed during the third replication. This resulted in randomly appearing step changes in altitude to the subjects who responded with large elevator deflections resulting increasing values of σ_{δ_e}, σ_θ and $\sigma_{\dot{\theta}}$.

Pilot ratings and pilot comments

Due to lack of experience with Cooper-Harper pilot rating scales, the briefing of the subjects appeared to have been lacking formal rigour. Apparently subjects had differing opinions on what exactly was to be rated.

Two subjects tended to base their ratings on both the pitch control of the simulated aircraft and the simulation of the pitching rotation. Nonetheless they turned out to be very consistent in their ratings and their comments were directed to the pitch control task as well as to the simulated motion. Of course this caused a difference in the mean and the standard deviation of the ratings between subjects. Still, significant differences in ratings due to different filters could be shown by analysis of variance, see Table 6. All ratings of each subject were normalized and the final values are presented in Fig. 11.

The lowest rating corresponds with the best performance being achieved with filter 5. No further evident correlation between pilot ratings and performance was found.

From the written comments, the following summary was composed.

Filter no. 1: Sensitive motion helps to control the pitch attitude.
Filter no. 2: Pitching rotation is sensitive but the additional information does not help much in the pitch control task.
Filter no. 3: Pitching motion response is too small and sluggish. Motion is confusing.
Filter no. 4: Motion confuses pitch control. Pitching motion response is sluggish and very small.
Filter no. 5: Controlability of pitch attitude is very good. Pitching motion response is direct. Motion improves pitch control.
Filter no. 6: Mostly no pitch response perceived. Lack of pitching motion does not confuse pitch control. Preferable to filters 3 and 4.

4. DISCUSSION

The optimization procedure yielded an unambiguous result only for the heave motion filter providing the normal accelerations.

It was mentioned in Chapter 2 that the simulation of the pitching acceleration and the specific forces along the X-axis can be done either by filtering θ and A_X or by filtering θ and $\dot{u}$. Results of any procedure to optimize either of these two ways of simulation, however, will depend on the form of the cost function and the weighting coefficients chosen. As long as the influence of the vestibular outputs on the perception process is not fully understood and incorporated in the cost function, results of optimization procedures will always be sub-obtimal.

By choosing a limited number from the many possible configurations and settings for the evaluation experiment, significant differences in subject performance and opinion due to motion filter characteristics were demonstrated. Best performance and most favourable opinion ratings were obtained by a one to one simulation of the angle of pitch, completely neglecting the influence of $\dot{u}$, resulting in a less accurate simulation of A_X.

It is very tempting to conclude from the experiment that accurate simulation of pitching accelerations (see Fig. 7e, filter no. 5) is far more important than accurate simulation of A_X (see Figs 7c and d, filters 3 and 4). Such a firm conclusion however, cannot be drawn yet.

It must be borne in mind firstly, that the time-histories of Fig. 6 are of manoeuvre motions only, whereas in the computer simulation and the piloted simulator experiment both manoeuvre motion and disturbance motion were present. Secondly it should be remarked that if $\dot{u}$ is completely neglected in the simulation (filter no. 5), the specific force A_{X_S} in the simulator is exactly equal to $g \sin \theta_a$. The possibility that a pilot may extract extra information on the angle of pitch from the specific force A_{X_D}, can, although no indication for this was obtained from the pilot comments, not be excluded. Care should be taken of course that a moving base simulated aircraft does not have better handling qualities than the aircraft in flight.

It is apparent that further research on the manner in which a human pilot processes information, in particular vestibular sensations, is needed before firm conclusions can be drawn. Moreover a similar piloted experiment for the case of manoeuvre motion only,appears worthwhile.

5. CONCLUSIONS

From the results and the discussion the following conclusions can be drawn:

1. Reaching a straightforward solution of the optimization of motion filters is not possible due to the lack of understanding of the perception process of the human pilot.

2. For the case where only pitch and heave motions are available to generate symmetrical motion cues and for the aircraft dynamics and control task used in the experiment described in this paper, far better pilot performance and opinion are obtained if the pitching rotations are simulated unfiltered (one to one) at the cost of completely neglecting the simulation of the horizontal accelerations.

3. Considering the present state of the art, motion filter optimization combined with an evaluation in a piloted simulator experiment, can lead to considerable motion simulation improvements.

6. REFERENCES

1. S.F. Schmidt and B. Conrad:
   Motion drive signals for piloted flight simulators.
   NASA CR-1601, 1970.

2. B. Conrad, S.F. Schmidt, J.G. Douvillier:
   Washout circuit design for multi degrees-of-freedom moving base simulators.
   AIAA Paper no. 73-923.

3. R.V. Parrish and D.J. Martin:
   Comparison of a linear and nonlinear washout for motion simulators utilizing objective and subjective data from CTOL transport landing approaches.
   NASA TN D-8157.

4. J. Gundry:
   Man and motion cues.
   Proceedings of the Royal Aeronautical Society, Third Flight Symposium on "Theory and Practice in Flight Simulation". 1976.

5. R.J.A.W. Hosman and J.C. v.d. Vaart:
   Vestibular models and thresholds of motion perception. Results of tests in a flight simulator.
   Delft University of Technology, Department of Aerospace Engineering, Report LR-265, 1978.

6. J.G. den Hollander and M. Baarspul:
   The flight simulator three degrees of freedom motion system of the Department of Aerospace Engineering of Delft University of Technology. Report LR-269, 1978.

7. L.R. Young:
Visually induced motion in flight simulation.
In: AGARD Conference Proceedings on Piloted Aircraft Environment Simulation Techniques, AGARD CP-249, 1978.

8. O.H. Gerlach:
The determination of stability derivatives and performance characteristics from dynamic manoeuvres.
Delft University of Technology, Department of Aerospace Engineering, Report VTH-163, 1971.

9. Anon.:
Approach and landing simulation.
AGARD Report No. 632, 1975.

10. R.L. Stapleford, S.J. Craig and J.A. Tennant:
Measurement of pilot describing functions in single-controller multi loop tasks.
NASA CR-1238, 1969.

ACKNOWLEDGEMENT

The authors wish to acknowledge the work done by mr. P. van Lent, student of aeronautical engineering, who carried out the optimization procedure and assisted in the design and execution of the simulator experiment and the cooperation of messrs. R.J. Kuil and P. de Lange of the University staff who participated as subjects.

TABLE 1. Aircraft geometric parameters and aerodynamic coefficients

W	= 2315 kg	V	= 52 m/sec
S	= 23,23 m^2	C_L	= 0.6
h	= 450 m	b	= 14.63 m
$\bar{c}$	= 1.5875 m	$2\mu_b$	= 11.12
$2\mu_c$	= 97.963	$2\mu_b K_x^2$	= 0.094
K_y^2	= 1.283	$2\mu_b K_z^2$	= 0.2135
$2\mu_c K_y^2$	= 125.686		

Aerodynamic coefficients

C_{X_o} = 0	C_{Z_o} = -0.673	C_{m_u} = +0.031
C_{X_u} = -0.18	C_{Z_u} = -1.27	C_{m_α} = -1.051
C_{X_α} = 0.374	C_{Z_α} = -5.74	$C_{m_{\dot{\alpha}}}$ = -2.817
	$C_{Z_{\dot{\alpha}}}$ = -0.74	C_{m_q} = -16.3
	C_{Z_q} = -4.32	C_{m_δ} = -2.0
	C_{Z_δ} = -0.407	

$C_{Y_\beta} = -0.684$	$C_{l_\beta} = -0.0805$	$C_{n_\beta} = 0.031$
$C_{Y_r} = -0.195$	$C_{l_p} = -0.6319$	$C_{n_p} = -0.038$
$C_{Y_{\delta_r}} = 0.0607$	$C_{l_r} = 0.0926$	$C_{n_r} = -0.057$
	$C_{l_{\delta_a}} = -0.18$	$C_{n_{\delta_a}} = -0.003$
	$C_{l_{\delta_r}} = 0.006$	$C_{n_{\delta_r}} = -0.062$

TABLE 2. Data for Dryden turbulence, see Ref. 8.

$\sigma_{u_g} = 1$ m/sec

$\sigma_{w_g} = 1$ m/sec

$L_{u_g} = 160$ m

$L_{w_g} = 160$ m

TABLE 3. Pilot model parameters.

Inner loop

$K_\theta = 1.0$

$\tau_L = 0.4$ sec

$\tau_I = 0.7$ sec

$\tau_c = 0.3$ sec

$\omega_c = 1.8$ rad/sec

$\varphi_m = 43^\circ$

Outer loop

$K_H = -0.019$

$\omega_c = 0.61$ rad/sec

$\varphi_m = 60^\circ$

TABLE 4. Single degree of freedom performance of the three degrees of freedom motion system of the flight simulator.

	Travel	Speed	Acceleration
Heave	$\pm$ 0.3 m	$\pm$ 1 m/sec	+ 15 m/sec^2 - 28 m/sec^2
Roll	$\pm 16^\circ$	$\pm 73^\circ$/sec	$\pm 1090^\circ$/sec^2
Pitch	$\pm 15.5^\circ$	$\pm 50^\circ$/sec	+ 315°/sec^2 - 460°/sec^2

TABLE 5. Mean values of Pilot Ratings and the standard deviations of the measured variables.

Filters	1	2	3	4	5	6
P.R.	3.9	5.23	5.97	5.2	2.53	5.9
P.R. normalized	-.508	+.253	+0.679	+.243	-1.320	+0.632
σ_θ °	1.354	1.348	1.245	1.426	0.987	1.221
$\sigma_{\dot\theta}$ °/sec	1.101	1.041	1.059	1.209	0.999	1.041
σ_h m	5.28	5.60	5.43	5.61	5.48	5.35
σ_{δ_e} °	0.370	0.347	0.352	0.418	0.327	0.354
σ_ψ °	1.989	1.920	1.908	2.126	1.960	1.891
σ_φ °	2.321	2.195	2.195	2.393	2.285	2.213
σ_{δ_a} °	1.248	1.169	1.169	1.306	1.238	1.193

TABLE 6. Results of the analyses of variance.

	P.R. normalized	σ_θ	$\sigma_{\dot\theta}$	σ_h	σ_{δ_e}	σ_ψ	σ_φ	σ_{δ_a}
Filters	**	**	**	-	**			
Subjects	-	-	**	**	**	-	**	**
Interaction Filters Subjects	-	-	-	-	-	-	-	-
Replications	-	-	**	*	**	-	-	**

* $\alpha < 0.05$

** $\alpha < 0.01$

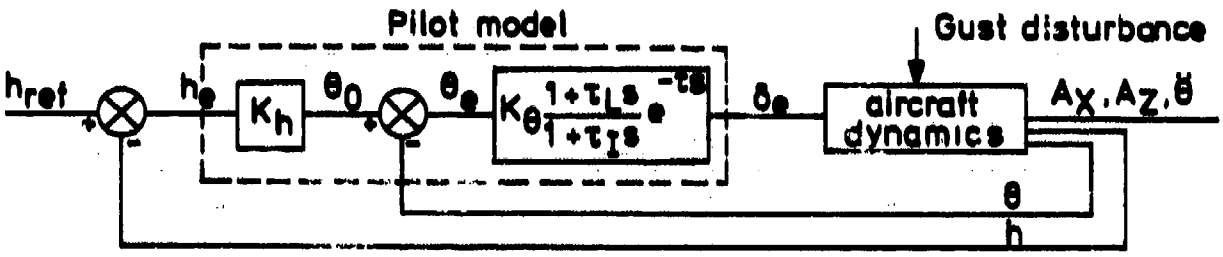


Fig. 1a : Simulated Pilot - Aircraft control loop

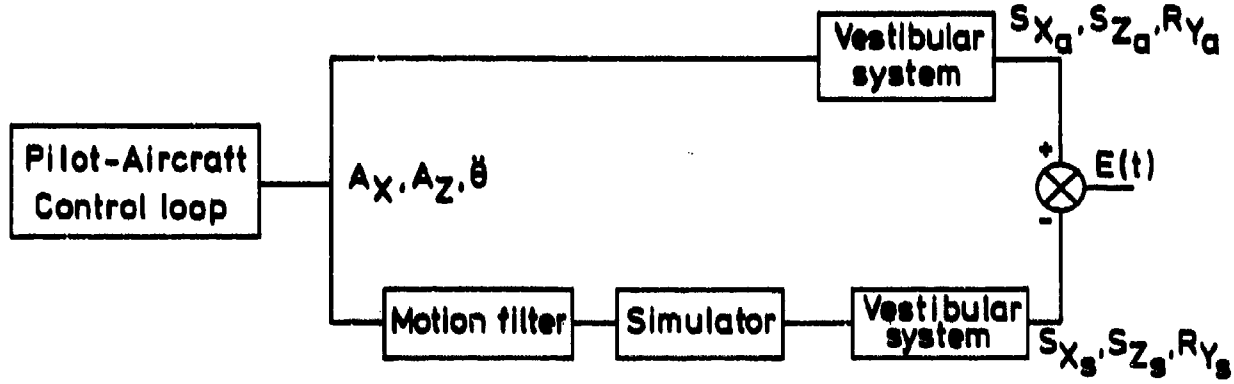


Fig. 1b : Computation scheme to determine the error E(t) of the vestibular output in the simulator.

$$E(t) = \begin{bmatrix} S_{X_a} - S_{X_s} \\ S_{Z_a} - S_{Z_s} \\ R_{Y_a} - R_{Y_s} \end{bmatrix}$$

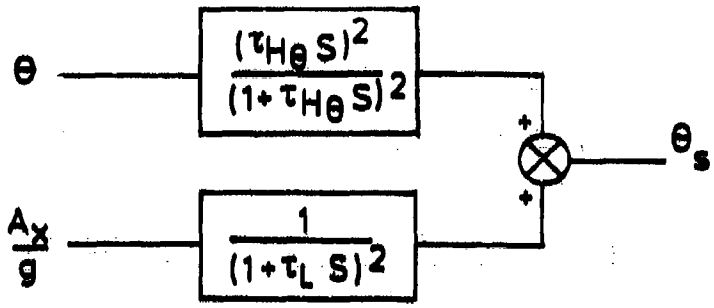


Fig. 2 : Filter configuration A for the control of the simulator pitch attitude.

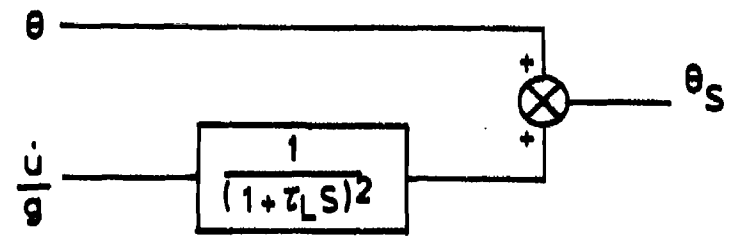


Fig. 3 : Filter configuration B for the control of the simulator pitch attitude.

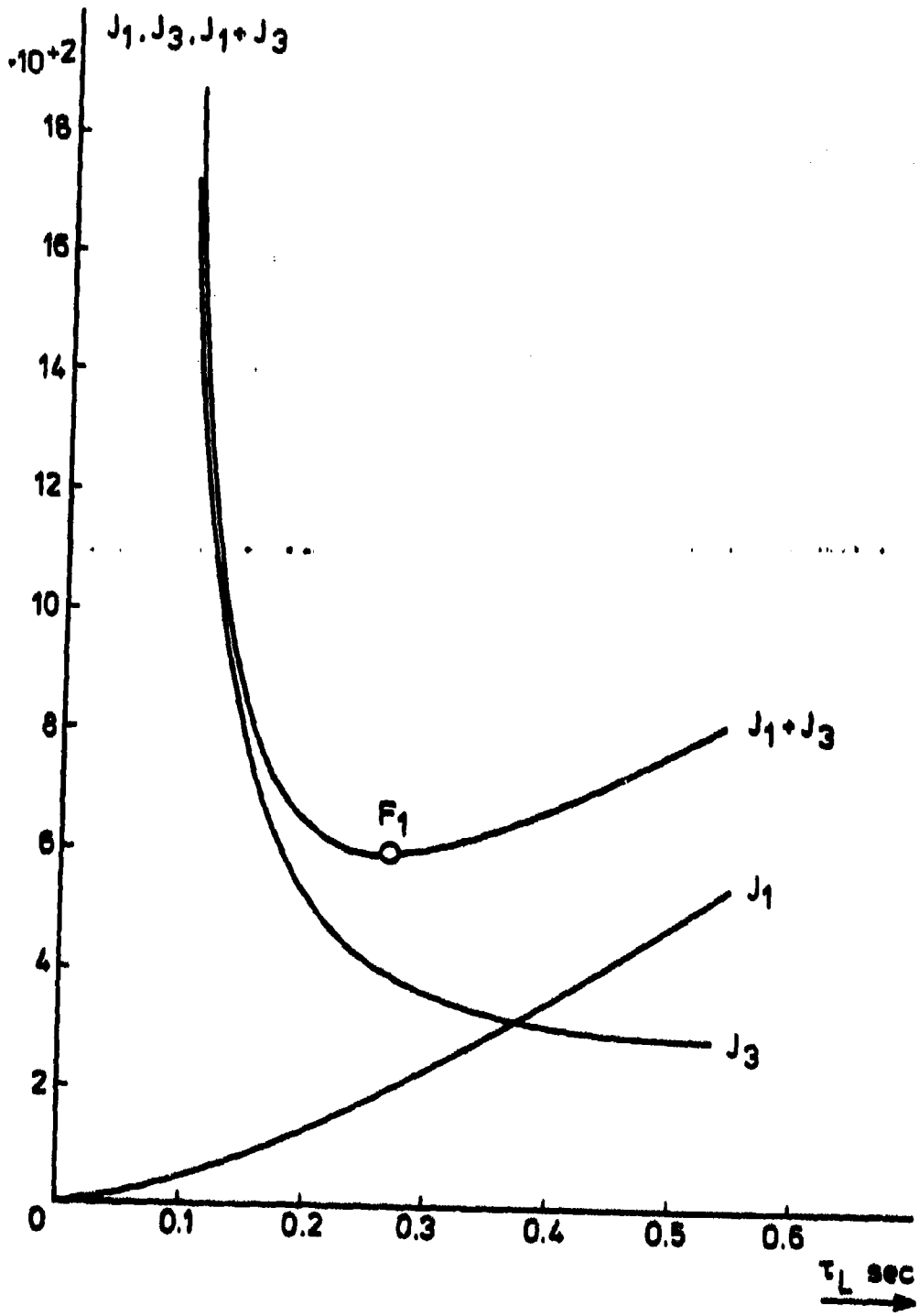


Fig. 4 : The influence of the low pass filter time constant τ_L in the filter B configuration on the cost function components J_1 and J_3.

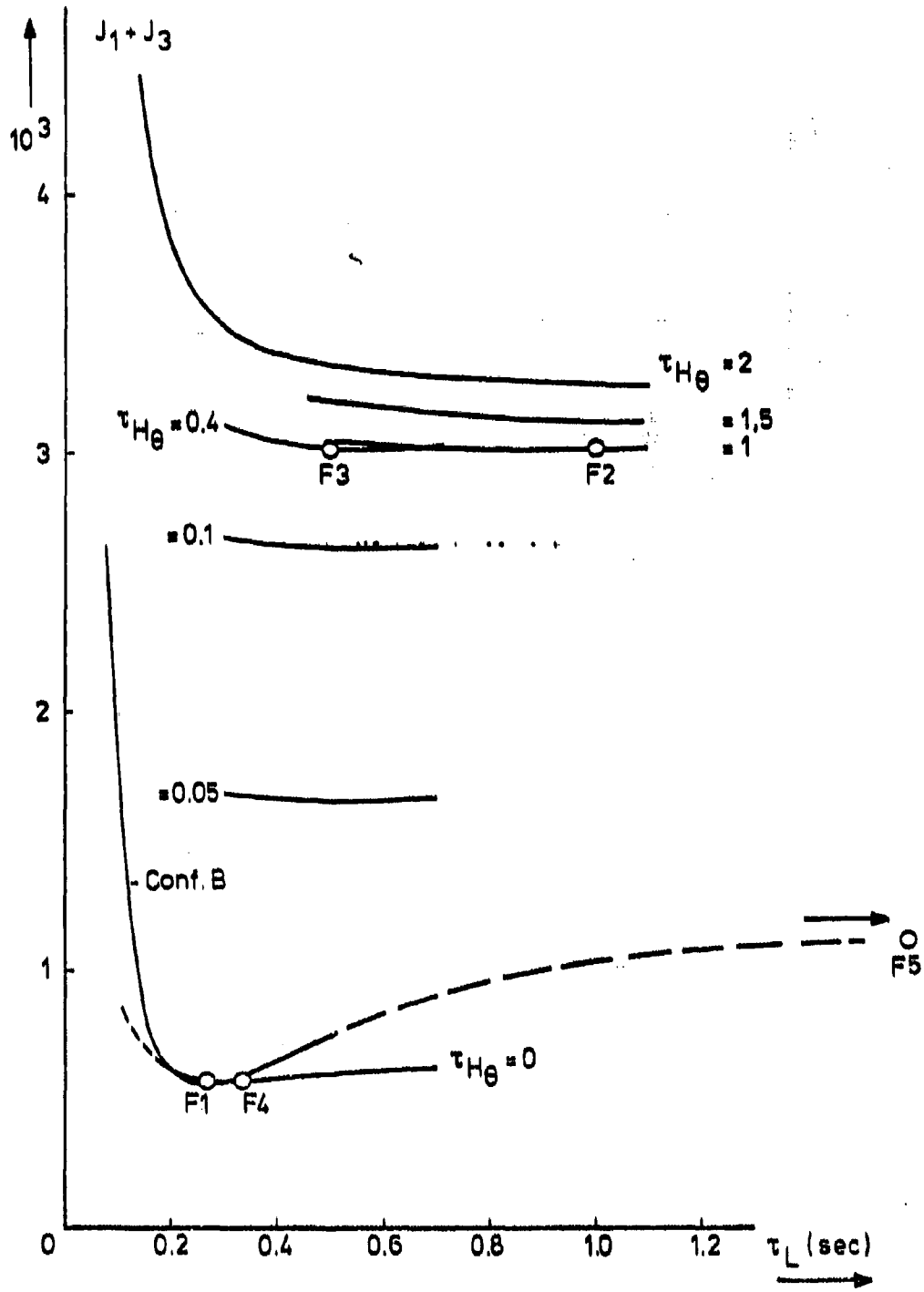


Fig. 5 : The influence of the time constants τ_{H_θ} and τ_L in the filter A configuration on the cost function component $J_1 + J_3$

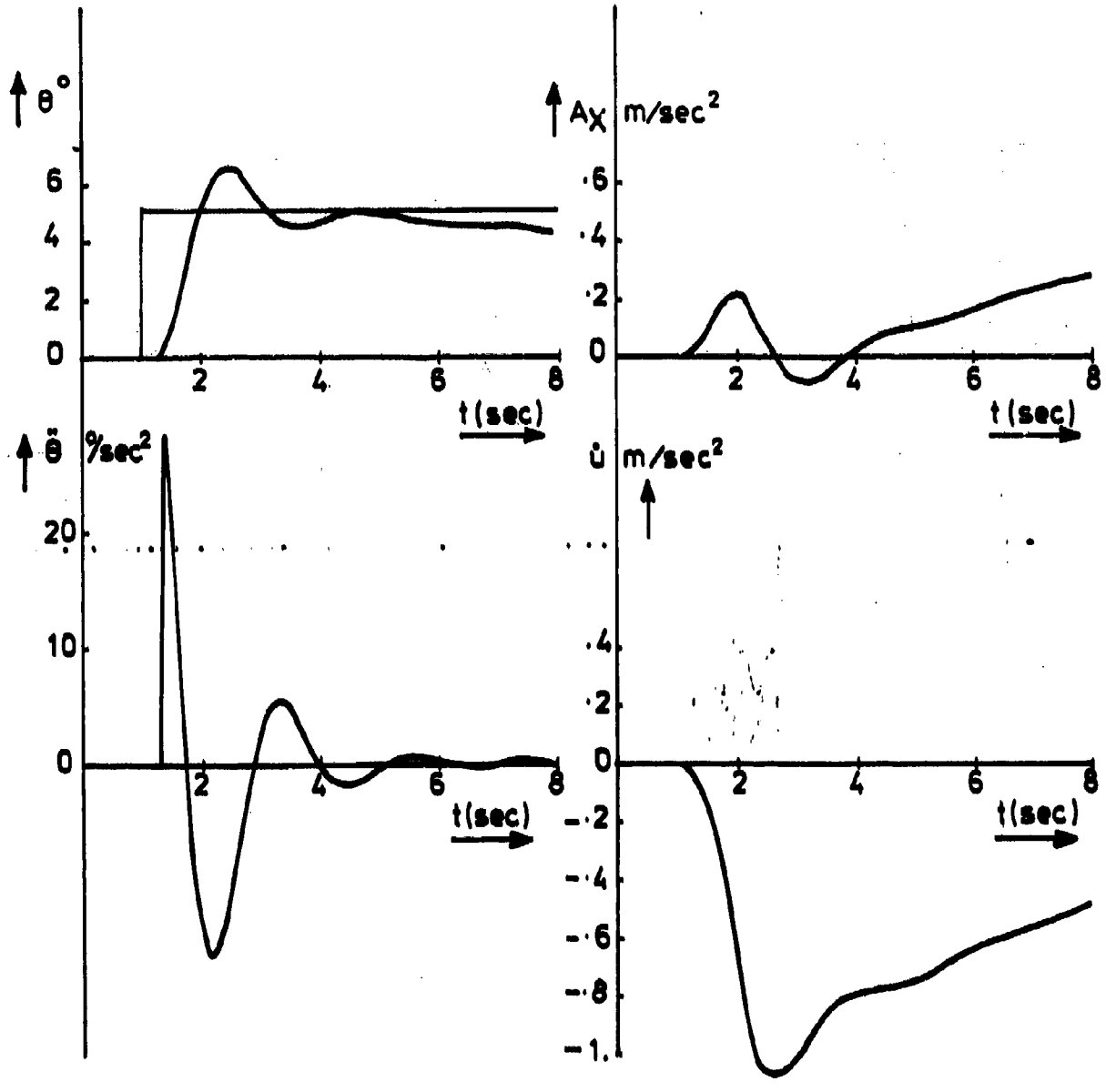


Fig. 6 : The responses of θ, θ̈, A_x and u̇ of the simulated aircraft on a commanded stepchange of θ

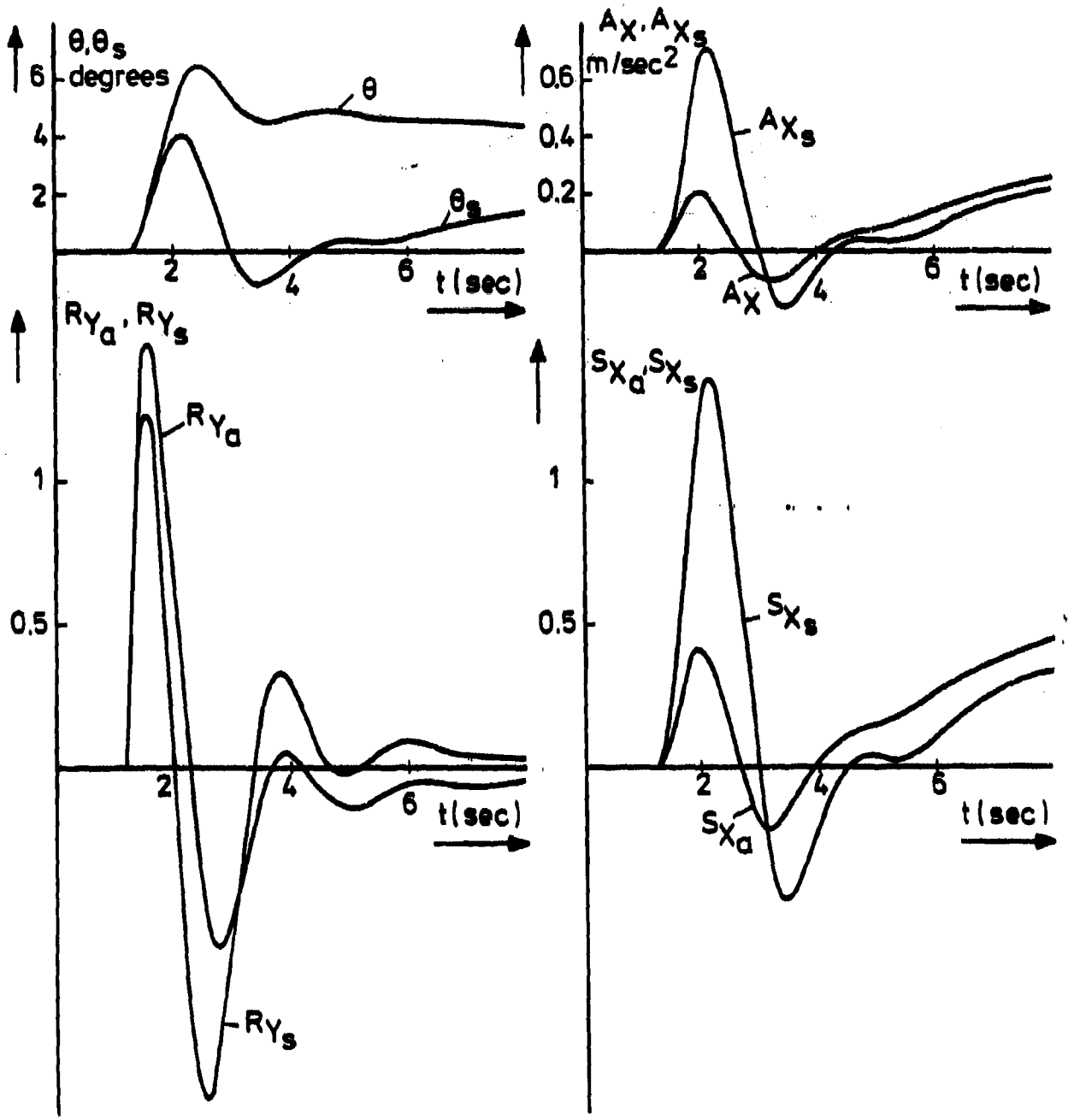


Fig. 7a : Comparison of the responses of aircraft and simulator and the corresponding vestibular responses due to Filter 1.

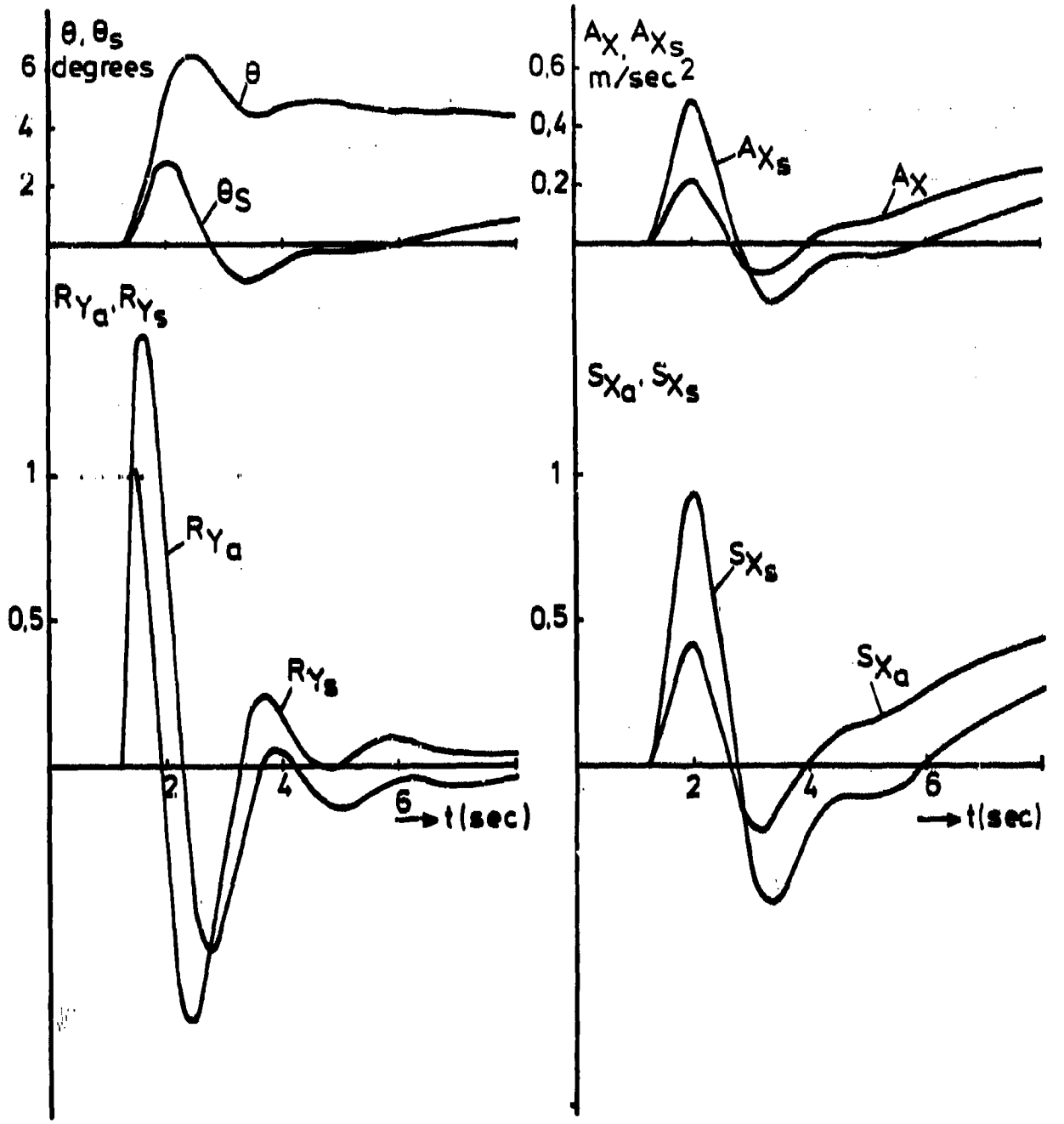


Fig. 7b: Comparison of the responses of aircraft and simulator and the corresponding vestibular responses due to Filter 2.

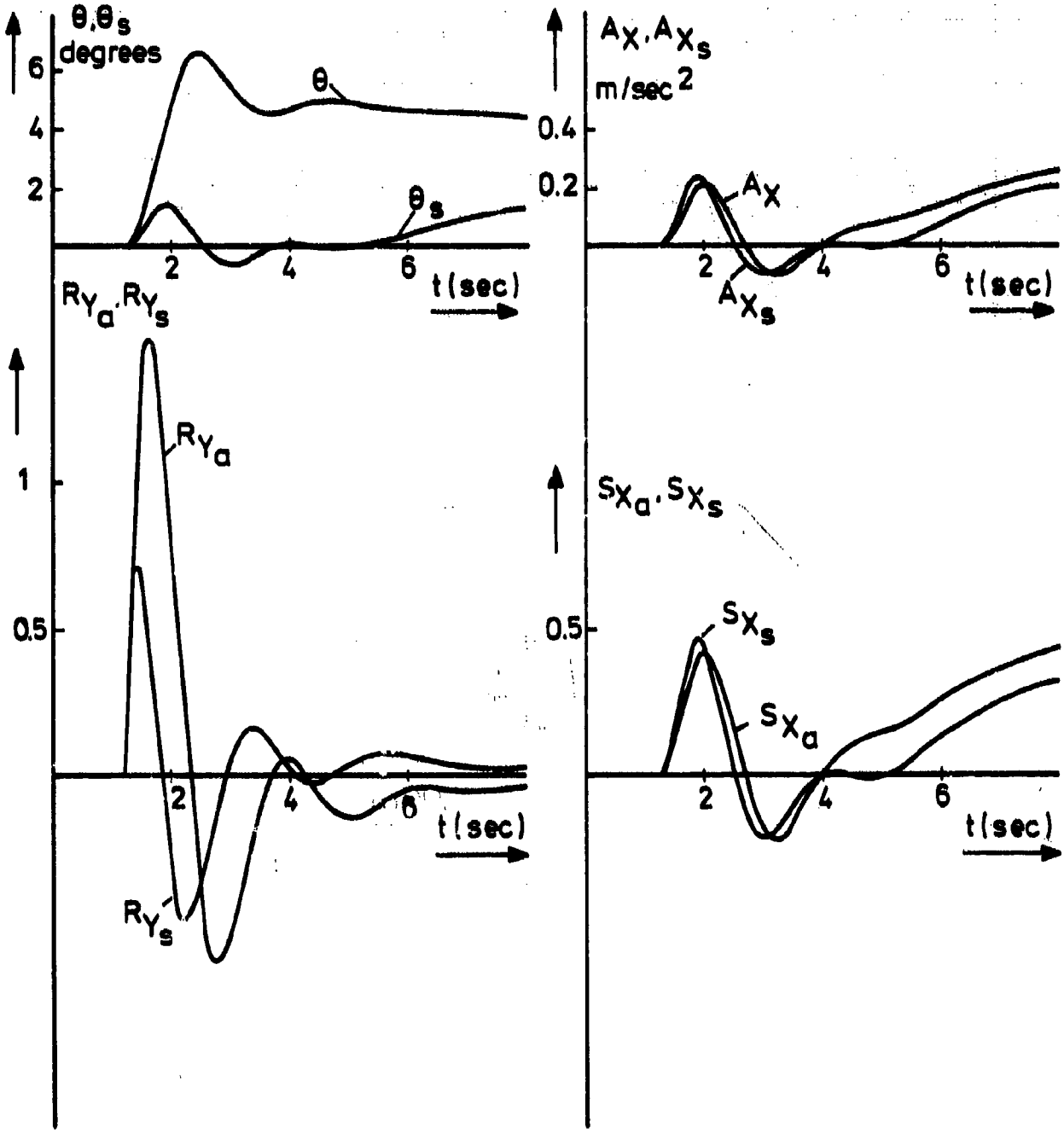


Fig. 7c : Comparison of the responses of aircraft and simulator and the corresponding vestibular responses due to Filter 3.

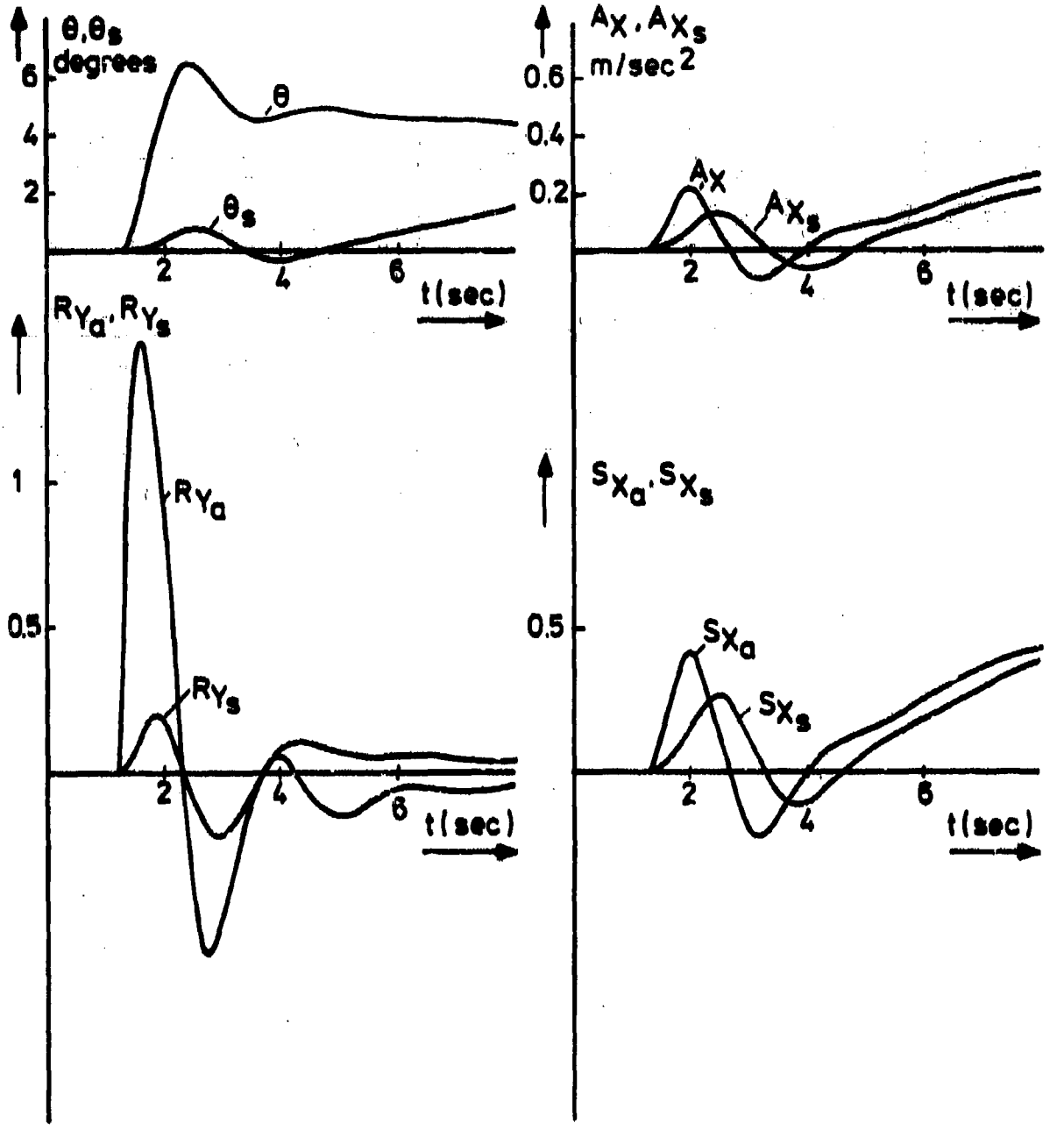


Fig. 7d: Comparison of the responses of aircraft and simulator and the corresponding vestibular responses due to Filter 4.

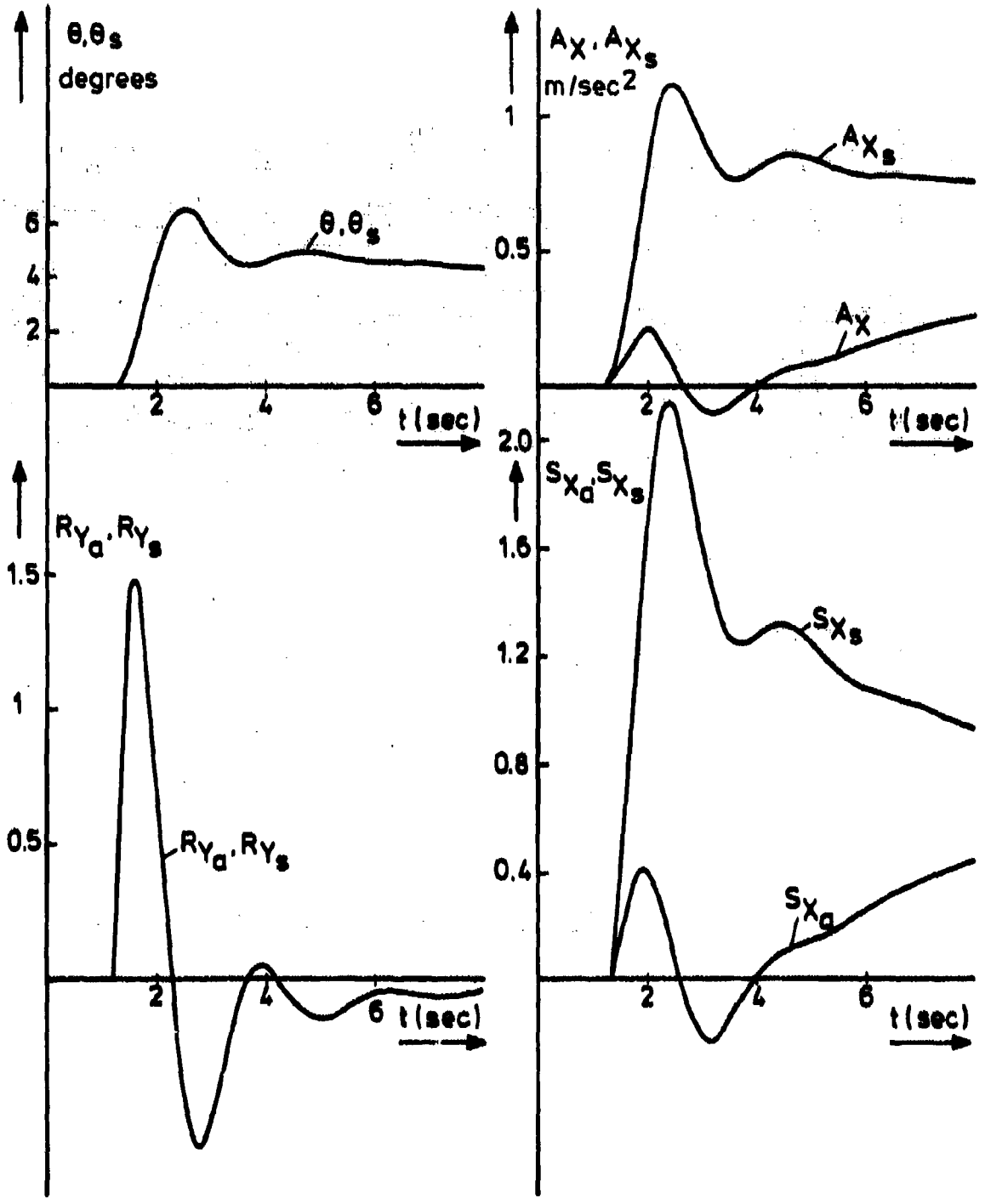


Fig. 7e: Comparison of the responses of aircraft and simulator and the corresponding vestibular responses due to Filter 5.

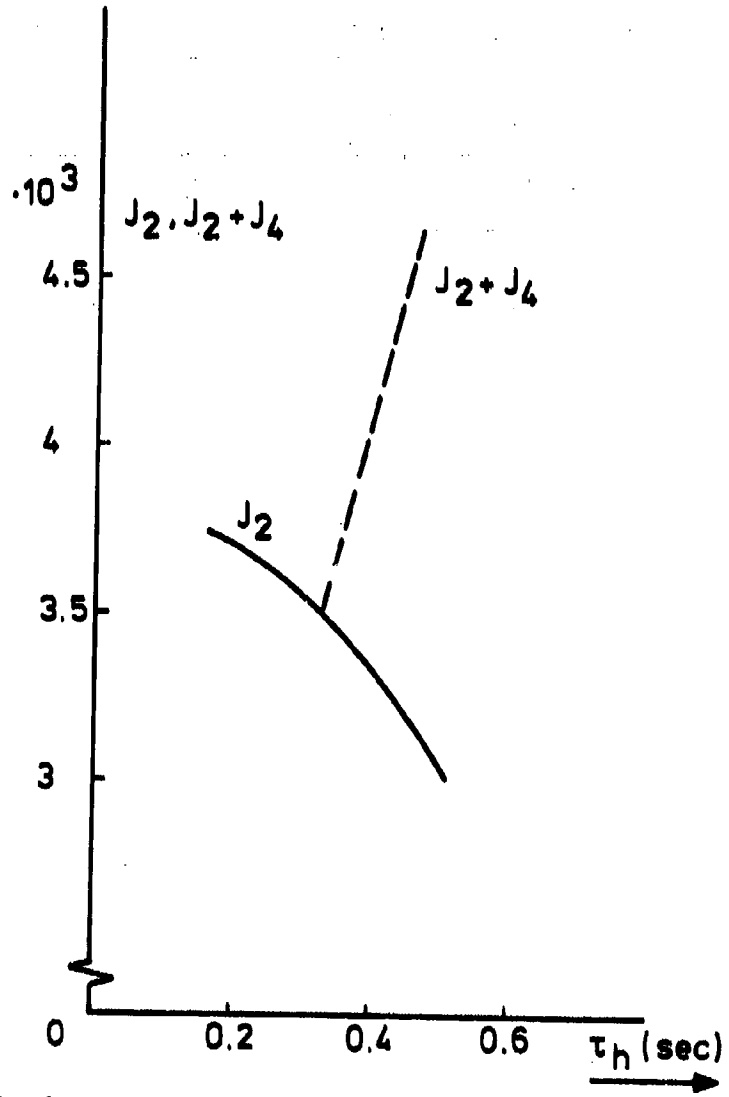


Fig. 8 : The influence of the time constant of the high pass A_Z filter on the cost function components J_2 and J_4.

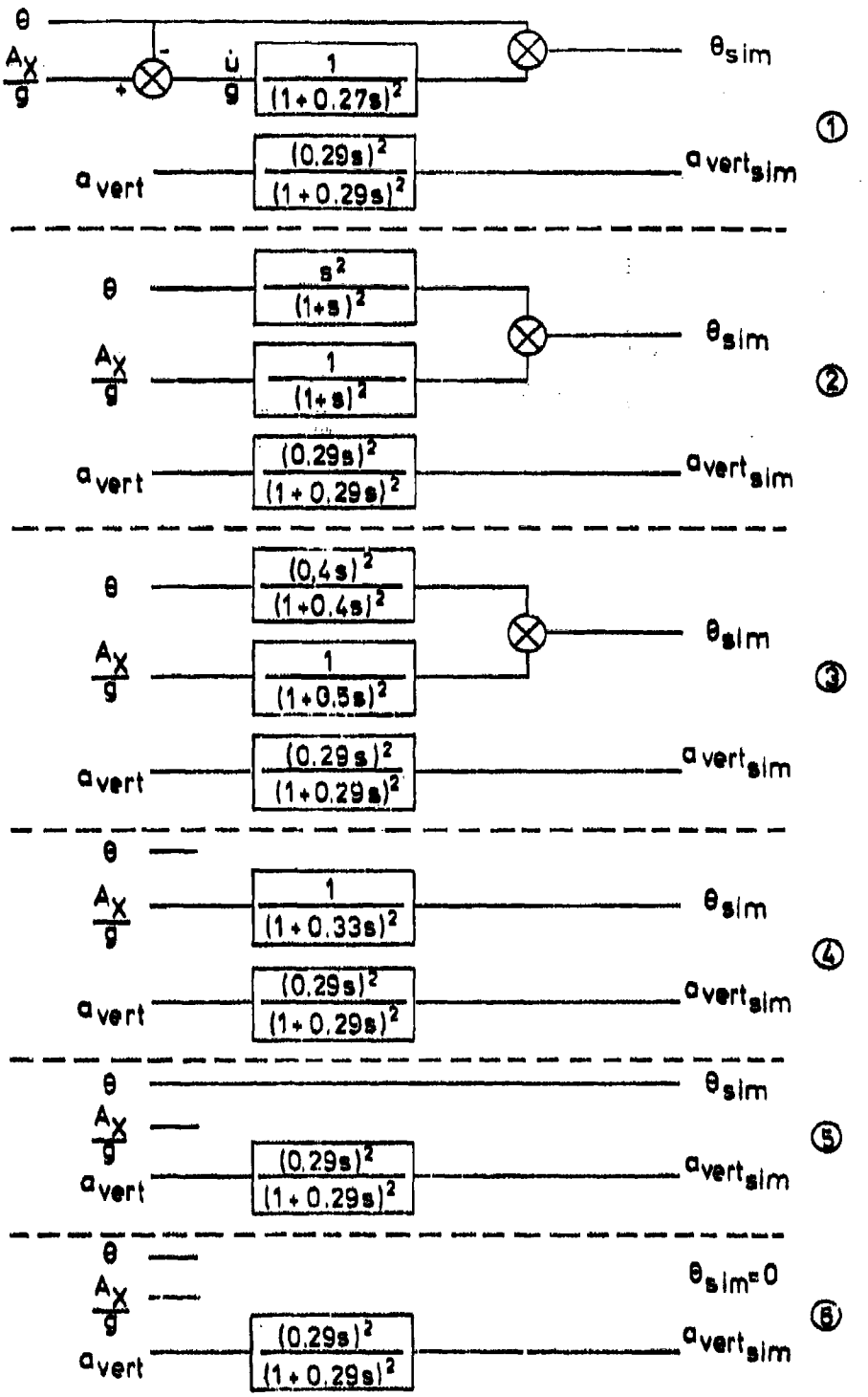


Fig. 9 : Filter combinations as used in the experiment.

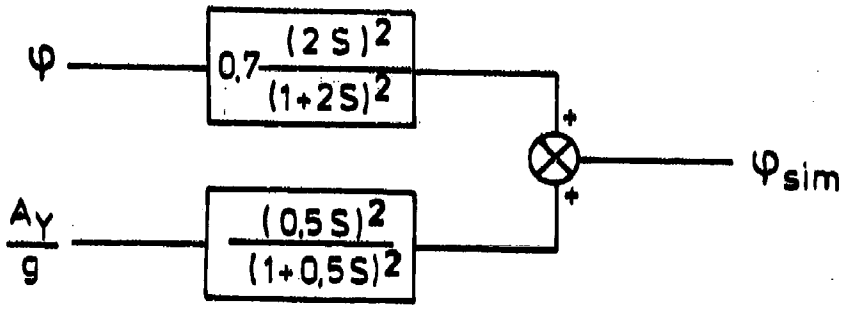


Fig.10: Motion filters for generation of asymmetric motion cues

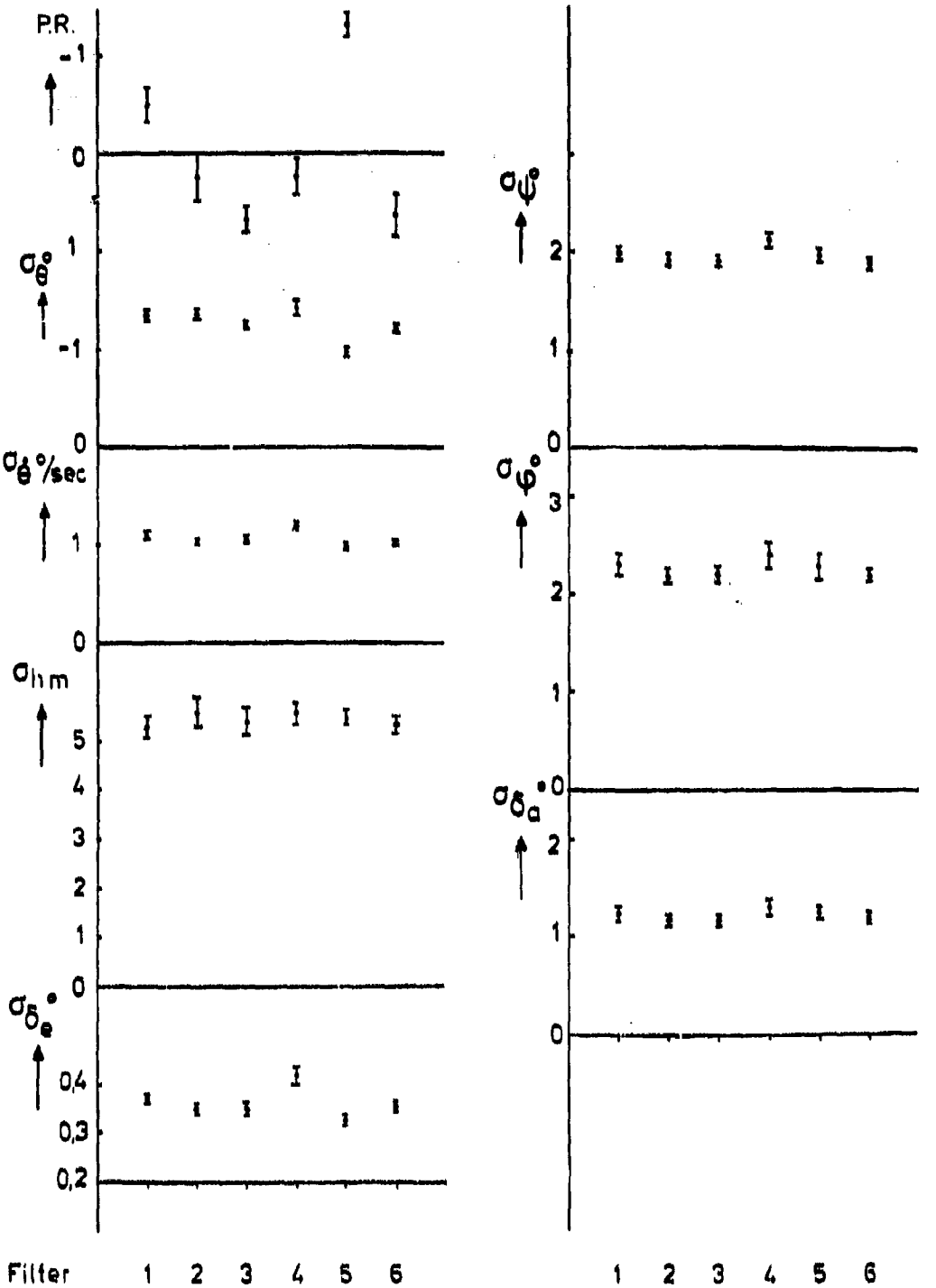


Fig.11 : Pilot ratings and performance data .
Mean values of 15 runs (3 subjects, 5 replications)

HUMAN PILOT PERCEPTION EXPERIMENTS

Dr. David L. Quam
Aerospace Engineering
University of Dayton
Dayton, Ohio 45469

ABSTRACT

This paper describes a set of experiments on the LAMARS simulator to measure human pilot perception of motion. Various combinations of visual field motion and platform motion are used to experimentally determine the minimum platform motion necessary to stimulate adequate perception of the visual field motion. An experiment is conducted in a single axis (such as yaw) while a tracking task is performed in a cross axis (such as pitch) to represent a typical flight workload. The subject's "gut feeling" motion perception is indicated in real time using the cockpit control for that axis (rudder).

EFFECTS OF VARIOUS LATERAL-BEAM-MOTION WASHOUTS ON PILOT TRACKING AND OPINION IN THE "LAMAR" SIMULATOR

by Henry R. Jex, Wayne F. Jewell, Raymond E. Magdaleno
Systems Technology, Inc., 13766 S. Hawthorne Blvd.
Hawthorne, CA 90250
and Andrew M. Junker, 6570th Aerospace Medical
Research Lab., WPAFB, OH 45433

ABSTRACT

A series of moving-base flight simulator experiments was performed using roll and sway motions of the Large Amplitude Multimode Aerospace Research Simulator (LAMARS) of the Flight Dynamics Lab at WPAFB, OH. The objectives were to:

a) Tie in the roll-only results of the 4 experienced pilots used here with previous results (Ref. 1) for 4 well-trained nonpilot subjects.

b) Investigate effects of various lateral-beam-motion "washout" filters designed to keep the lateral sway within the ±10 ft of LAMARS travel.

The high-pass washouts on lateral beam travel (y_{beam}) were of the general form:

$$\frac{y_{beam}}{y_{free\ flight}} = \frac{K_y\, s^2}{s^2 + 2\zeta_y \omega_y s + \omega_y^2}$$

where: K_y = attentuation factor, ω_y = high-pass break frequency (r/s), and ζ_y = 0.70 (fixed).

A range of K_y and ω_y was explored, from which example data are shown. A nonlinear (time varying) washout was also tested in which ω_y was continuously adjusted so as to permit correct cues for small roll activity, while reducing the travel peaks for large roll angles. Reshaping the forcing functions was also investigated.

The basic task was to follow an evasive (randomly rolling) target while suppressing gust disturbances (Ref. 1). A two-independent-input technique produced behavioral data (describing functions) and performance data (error and control scores), which revealed how pilots used the visual and motion cues. Subjective data was also gathered on the tracking task as well as on limited "sidestep" maneuvers.

The main results show: excellent tie-in with prior roll-only experiments with non-pilot subjects; most tracking performance and behavioral parameters were not significantly affected by various degrees of sway washout; pilot commentary became more consistent and adverse as the spurious side-force peaks exceeded about 0.1 G_y; specific problems were mapped vs. ω_y and K_y.

THE PROBLEM

Before describing the specific objectives of this experiment, let us review the problem of simulating motions of aircraft involved in lateral maneuvers. Figure 1 shows some of the motion quantities involved in an idealized "turn entry" maneuver, wherein the pilot puts in a one second aileron pulse resulting in a 15 degree bank-and-stop maneuver. If the aircraft were well coordinated there would be no lateral-specific-force (LSF = apparent sideways acceleration on the pilot) i.e., the apparent gravity vector would always parallel the pilot's spine Vehicle acceleration caused by the banking of the lift vector would be nulled out by the aircraft accelerating to the right. The motions for free flight are shown by the dotted lines in Fig. 1. The main problem for simulation is evident in the bottom of Fig. 1; the simulator travel required to perfectly replicate free flight becomes excessive after only a few seconds.

To alleviate this problem two types of simulator motion "washout" are employed. First, the roll angle itself is gradually reduced to zero ("washed-out") from the otherwise constant value that would be present in free flight. This is shown by the second plot in Fig. 1, where a first order washout filter reduces the bank angle from its peak with a time constant of $T_\varphi = 1/0.4 = 2.5$ seconds. The resulting roll rate (the primary cue sensed by the pilot's vestibular system) is shown at the top. It can be seen that the rectangular pulse of the free flight case is distorted by this roll washout; it slightly reduces the main pulse and produces a reversal in roll rate ("rate miscue").

In a roll-only simulator where no lateral travel of the cockpit is permitted the pilot will also feel a false "tilt cue" given by: $a_y' \triangleq g \sin \varphi$. Washing out the roll angle reduces this false tilt cue.

Even with roll washout there remains some residual tile-cue as shown by the dashed line in the third plot of Fig. 1. This lateral-specific force can be further reduced by allowing the cab to accelerate laterally. However, even with the roll angle returning back to zero, as shown, the lateral travel required to remove all the LSF is excessive. Typically, the lateral motion (sway) drive also employs a first or second-order washout to bring the cab back towards the center of the travel limits, amount as shown at bottom of Fig. 1. This limits the motion to just under 10 ft laterally, and results in the LSF shown by the solid line in the plot above it. Two affects of sway washout are obvious from the a_y' plot: a) The peak miscue is reduced, as intended. b) The peak of the a_y' miscue is delayed relative to the pure-tilt cue by t_p seconds. These effects should be borne in mind when reading the rest of this paper; for example, it turns out that the advantage gained by reducing a_y' can be nullified by excessive delays with respect to the roll angle.

Even for this small maneuver, and with roll washout, the amounts of travel are excessive and require large sway washouts to remain within the simulator limits. The sway washout does not reduce the miscue significantly over the roll-only case. Extremely large values of lateral travel (on the order of 50 to 100 ft) would have to be provided to reduce the LSF miscue to under 0.1 g for typical bank and stop maneuvers. Thus, the key problems to be addressed here are: What is the "least-worst" sway washout which will a) produce behavior as similar as possible to the free flight case and b) produce miscues which are small enough to be acceptable by the pilots without seriously impairing the realism of the task.

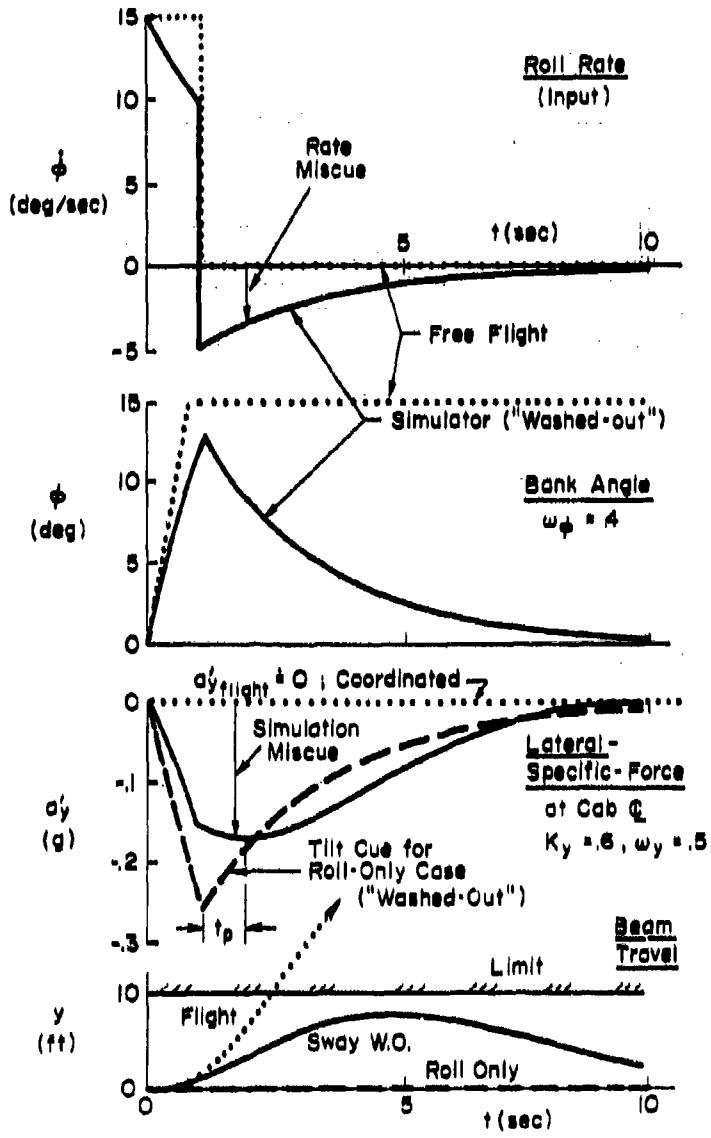


Figure 1. Comparison of an Idealized Turn-Entry in Free-Flight and a Motion Simulator Having "Washed-Out" Roll and Lateral Motions

Objectives

These experiments were a continuation of basic research done previously for the U.S. Air Force Aerospace Medical Research Labs on the Dynamic Environment Simulator (DES; a roll-only simulator)(Ref. 1). The current experiments were performed on the USAF Flight Dynamic Laboratories' "Large Amplitude Multimode Aerospace Research Simulator" (LAMARS), at WPAFB, Ohio. The LAMARS has five degrees of freedom; heave, sway, roll and (limited) yaw and pitch angles are provided to a one-man cab. The cab mounted on a 20 ft cantilevered beam can move vertically or horizontally through ±10 ft with compensating yaw and pitch motions such that it remains straight and level. High performance servodrives are provided ("flat" bandwidth to roughly 3 to 4 Hz). An overview of the development of the LAMARS is described in Ref. 2.

The current experiments extended the roll-only washout experiments performed on the DES to the case of both roll and "sway" motions performed on the LAMARS. The specific objectives included the following:

1. Perform tie-in runs to cross-validate the DES and LAMARS facilities, using roll-only cases.

2. Compare experienced military pilots vs. the well-trained non-pilots used in Ref. 1.

3. Investigate the effects of various types of sway washout.

4. Investigate a non-linear washout concept which showed promise in covering a large range of situations with one mechanization.

5. Correlate (to the extent possible) both objective effects on performance and piloting behavior with subjective evaluations of the motion cues felt by the pilots.

The results given in this paper cover the above objectives roughly in the order given above.

Experimental Set Up

Block Diagram

The basic experimental task procedures and variables were similar to those used in the roll-only experiments of Ref. 1, to which the reader is referred for background. The block diagram applicable to these LAMARS experiments is shown in Fig. 2, along with the relevant transfer functions.

The basic task scenario was that of roll tracking in an air to air gunnery type encounter, wherein the target aircraft was performing evasive maneuvers which the pilot tried to track by matching his roll angle with that of the target. Simultaneously, there were unseen roll disturbances applied to the aircraft, allegedly due to "encounters with the tip vortices of the target aircraft." Sums-of-seven (or eight) sinusoids, with amplitudes selected to approximate realistic target and disturbance spectra, were used for each quasi-random forcing function. The frequencies of the target and disturbance sinusoids were interleaved, such that disturbance input was completely uncorrelated with the target input. The provision of two

such independent inputs enables both the visual and motion dynamics of the operator to be determined, as documented in Ref. 1.

Figure 2 shows that the task bank angle resulting from the pilots control was subtracted from the target motions to produce a "compensatory" tracking error on the display scope. This was given by a small aircraft like symbol which only rolled. The task bank angle was washed out by a first-order filter to produce a bank angle command to the LAMARS roll drive.

When the lateral drive is fixed, the LSF consists only of the "tilt cue" as shown in the bottom half of the block diagram. In perfectly simulated free flight ($W_Y(s) = 1.0$) the beam lateral acceleration would exactly cancel the tilt cue, producing zero a_y. The sway washout filter $W_Y(s)$ is set up as a high pass filter to pass the high frequency motions thought to be of importance for motion cues, while reducing the longer term accelerations which integrate to form large beam travels. The accelerations of the modified lateral motions now provides imperfect cancellation of the LSF tilt cue. Also shown in the block diagram is a point of entry for other vehicle side forces which could be represented by the beam accelerations (e.g. due to rudder action) but these were not employed in this experiment.

Controlled Elements

Transfer functions of various blocks in the diagram are shown below in Fig. 2*. The controlled element was characteristic of a fighter having a neutrally stable spiral mode, a roll subsidence mode at 1.7 rad/sec, and an actuator mode of 5 rad/sec. When the DES simulator was used earlier, a simulator-drive mode having a frequency of 10 rad/sec and damping-ratio of 0.37 was unavoidably present and this is shown in the first controlled element denoted "DES". A second set of tests was made with these low-damped DES modes removed but with the roll subsidence modified to preserve roughly the same low frequency phase characteristics. This is denoted by the "smoothed" control element in Fig. 2. The former was used to tie in the results with the previous DES experiments while the letter was used to investigate the effects of the lateral beam washout.

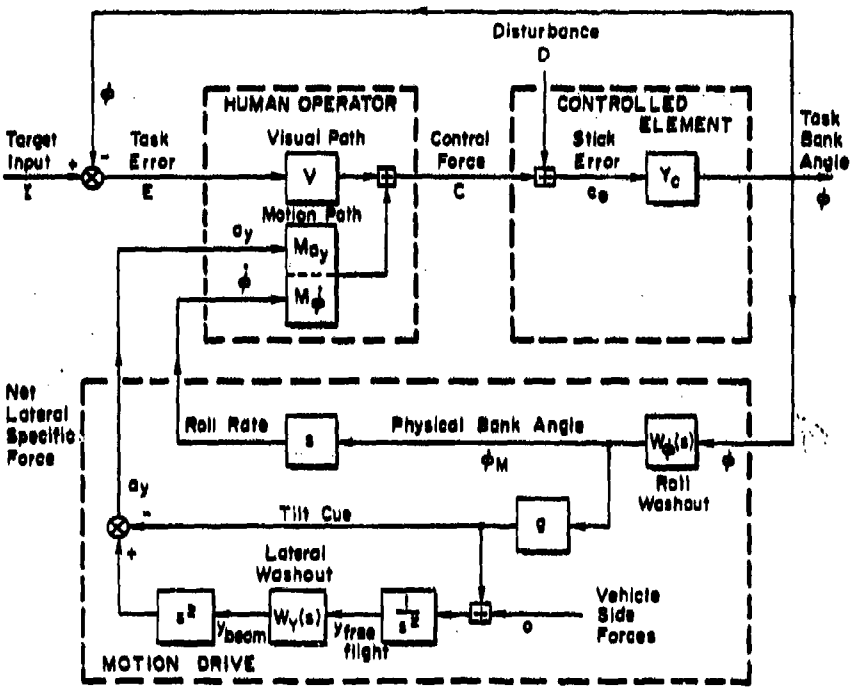
Washouts

To tie in with the prior DES experiments a number of roll-only runs were made which had no roll-washout. However, because residual tilts caused drifting of the beam under minimum beam-washout conditions, some roll washout was necessary for the remainder of the runs. Based on previous results given in Ref. 1, we selected a 0.4 rad/sec first-order filter roll washout.

As shown at the bottom of Fig. 2 the lateral beam washout was a second order high-pass filter; i.e., constant acceleration inputs resulted in a constant position output, and a constant rate input resulted in a centered output of the beam. The break frequency ω_y was varied from the minimum washout condition of about 0.2 rad/sec up to and above 1.0 rad/sec, which approaches the bandwidth of the closed-loop maneuvers. The damping ratio was held constant at 0.7. The washout filter also had an attenuation factor, K_y, which reduced the high pass level of the beam accelerations from 1.0 down to 0.20. The various combinations of K_y and ω_y will be given later.

---

*In our shorthand notation for transfer functions, pure numbers are gains, numbers in (parenthesis) are first order break frequencies, and numbers in [square brackets] are damping-ratio and frequency, respectively.



Controlled Elements:

$$Y_c(s) = \frac{\phi}{c} = \frac{\overset{\text{Gain}}{15000.}}{\underset{(1/T)}{(0.)}\underset{}{(1.7)}\underset{}{(5.0)}\underset{[\zeta,\ \omega]}{[.37,\ 10.]}} \quad \frac{\text{deg}}{\text{lb}} \equiv \text{"DES"}$$

or

$$= \frac{100.}{(0.)(1.0)(5.0)} \quad \frac{\text{deg}}{\text{lb}} \equiv \text{"Smoothed"}$$

Roll Washout: (First-order)

$$W_\phi(s) = \frac{\phi_M}{\phi} = \frac{1.0(0)}{(.4)}$$

Lateral Beam Washout: (Second-order)

$$W_Y(s) = \frac{Y_{Beam}}{Y_{free\text{-}flt.}} = \frac{\cancel{K_Y}(0)(0)}{[.7,\ \cancel{\omega_Y}]} \quad \longleftarrow \text{Main Variables}$$

Figure 2. Block Diagram and Transfer Functions

In another set of runs a non-linear mechanization was employed, which varied ω_y as a complex function of the roll angles and computed beam states, such as to maintain the minimum degree of washout consistent with the anticipated lateral motions. Further details on the nonlinear sway washout are given in Ref. 3.

Operations

The LAMAR3 motion simulator is a highly automated facility with pre-programmed, digitally-controlled parameters, designed for high-production simulation of systems. However, in current development these preset-able features which make it efficient for high volume usage, also make it difficult to explore "on-line" a range of variables in which process the experimenter varies parameters while the test is proceeding. Through the concentrated efforts of John Bankovskis of the Flight Dynamics Lab, the LAMARS was temporarily reconfigured for on-line parameter exploration. Many of the parameters which are shown in later tables as fixed variables in the experiment were first explored by these techniques; i.e. K_y and ω_y were varied in systematic and random manners, to map out the boundaries of subjectively noticeable effects. Such data are highly variable, interactive, and are very difficult to average and to present in a systematic manner. Nevertheless, much of the effort of the simulation had to go into these exploratory runs, in order to whittle down the number of parameters to those formally tested.

Among the objective data recorded were: the task roll errors and control activities and various motion quantities. In addition, signals were recorded at various points in the loop from which the pilot describing functions in response to both visual and motion inputs could be computed. The data reduction program was developed by AMRL from earlier work by BBN (Ref. 4) with dual-input additions by STI (Ref. 1). Adding sway measures and the transfer from AMRL to AFFDL computers was part of this project.

Using these reduced describing functions, STI performed the data averaging and model fitting and parameter identification using the techniques and Model Fitting Program (MFP) described in Ref. 1.

Subjects and Procedures

Four experienced military pilots were used as subjects. Three of the four had experience as instructors in various types of military aircraft and all had time both in fighters and heavier aircraft. Although these pilot subjects were not trained test pilots, they brough extensive experience in the real flight situation to the simulation, an overview which was missing in the non-pilots tested previously in Ref. 1. Therefore, these pilots could give commentary on the realism of each simulation case. We were also interested to see if they adopted the same or different tracking behavior as the previous non-pilots.

All the pilots were given several sessions to train on the LAMARS simulator, in the various tasks. It was found that the experienced pilots learned the task within a few sessions and achieved a fairly stable performance within a few dozen runs. There was a dramatic reduction from the training regimen which had to be used with the previous non-pilots (where hundreds of runs over a few weeks period were required to train them to asymptotic performance). As will be shown, the pilots and non-pilots adopted the same strategies, so the reduction in training time was due to the pilot's flight experience.

Most of the formal washout runs were made with random ordering of the combination of $k_{\dot{y}}$ and ω_y, and then pilot comments were solicited as to any noticeable effects of the simulation when compared with flight. We attempted a motion rating scale evaluation, but this was not as informative as the comments themselves.

The operating procedures were generally as follows: The pilot was allowed a warmup run and then did the roll-only (tie-in with DES) runs. Then the pilot was led through the various degrees of lateral beam washout. If the pilot could accept the range of washouts that had been previously established for other pilots, he was then run through a more formal set of runs in which describing function data were recorded, along with his subjective comments. In almost every case two or more were possible at each condition. The data were so consistent that further runs were felt to be unnecessary; instead, the nonlinear washout runs were inserted as a third session for each pilot.

Experimental Design

Table 1 shows the various combinations of control element, input and sway washout used in these LAMARS experiments. The "DES" and "smoothed" control elements have been described earlier. The "reference input" was identical to that used in Ref. 1 and contains large-amplitude low-frequency roll motions which elicit significant tilt cues in the roll-only simulation. Cancelling these roll angles required large lateral travel of the beam, so that a high amount of lateral beam washout had to be used in order to keep the beam within limits. Consequently, a "modified" input was employed, in which the low frequencies were attenuated to reduce the lateral travel requirements at any given level of beam washout. In effect, the input was washed-out instead of the output. This was fairly successful in reducing beam travel requirements without significantly affecting the pilot's control behavior near crossover frequencies. However, this also reduced the level of tilt angles so that tilt cues were not as subjectively apparent to the pilot as in the reference input case.

The amount of sway washout shown varies from: "none" (typical of free flight) through "low," "medium," and "nonlinear" (similar to the medium washout), and "high" washout. "High" washout is only slightly different from a roll-only case (no sway motion) while for the "static" case there is no motion at all. The trade-off between travel requirements and miscues is noted at the bottom of Table 1. The several variables mentioned above could result in an excessive number of test combinations but this was avoided by making selected comparisons, as shown by the circled "cases" shown in Table 1. In effect, these comparisons investigated these sub-experiments within the larger plan.

1. Tie in with prior DES experiments. Case (1) was run with identical control elements to the prior DES experiments, only difference being the use of experienced pilots.

2. Controlled element effects. Case (9) was run with the "smoothed" control element more typical of real aircraft to compare its results with the somewhat atypical DES dynamics.

3. Sway effects. Comparisons of cases (7) and (9) show the effects of sway motion with a medium washout, as required to maintain lateral sway travel within limits. It was

TABLE 1. EXPERIMENT DESIGN

PLANT Y_c	INPUTS Φ_{TT}, Φ_{DD}	ROLL WASHOUT ω_ϕ	AMOUNT OF SWAY WASHOUT						
			NONE	LOW	MED.	N.L.	HIGH	ROLL ONLY	STATIC
			$K_y = 1.0$ $\omega_y = 0$ "FREE FLIGHT"	1.0 .25	.6 .5	.6 .2↔.8 NON-LINEAR	1.0 1.0	1.0 ∞	
"DES"	"REFERENCE"	0.0						Tie-in w/D.E.S. ① Y_c	
"SMOOTHED"	"REFERENCE"	0.4	σ_y = 160 ft		Nonlinear W.O. Case ⑦ ↔ 8C; Sway Effects			Effects ⑨	
"SMOOTHED"	"MODIFIED" (Reduced at low frequencies)	0.4	σ_y > 40 ft	③	Input Effects ⑥	⑤		28.4	28S

Washout Amount

← MORE MIS-CUE →

← MORE TRAVEL →

not possible to compare a "no-washout" case because the lateral travel would have to exceed 160 ft.

4. Input Effects. To allow a lower degree of sway washout the "modified" input described earlier was used. This reduced the idealized free flight motion from 160 down to 40 ft and allowed a smaller degree of washout to maintain the cab within simulator limits. Comparing configuration (6) with (7) documented the effects of this change in input alone. The resulting effects of input shaping on the describing functions were negligible and this comparison will not be further discussed in this paper.

5. Sway Washout Variations. These were the primary runs of the experiment. With the "modified input" a range of increasingly severe washouts could be used, corresponding to cases (5), (6), (5), and (8.4), respectively.

6. Non-Linear Washout. Comparison of cases (7) and (8c) allowed us to determine the effects of the non-linear washout on performance and subjective ratings.

EXPERIMENTAL RESULTS AND COMPARISONS

Tie-In With DES Experiments

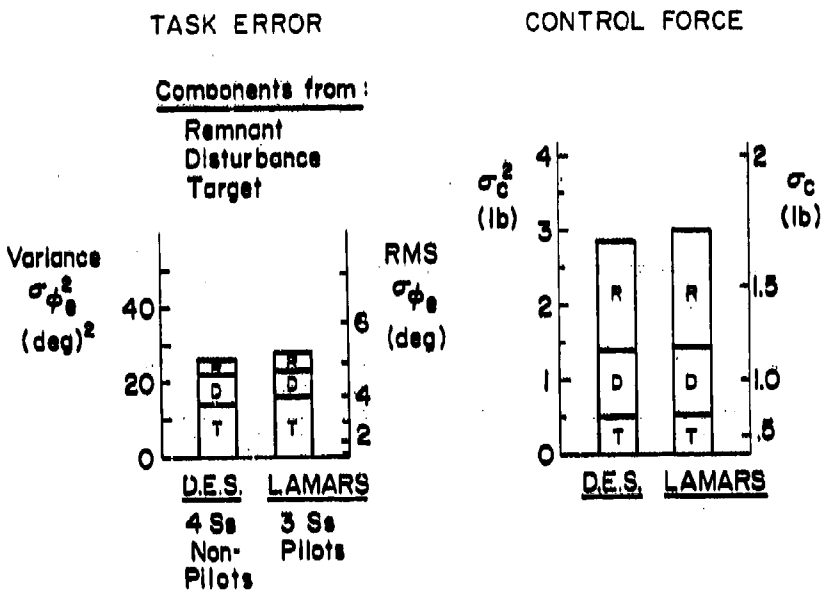
The first goal was to establish the validity of a LAMARS simulation of the previous DES experiments for the roll only case. The same display, control stick, controlled element and drive-logic dynamics were used, as validated by describing function measurements. Figure 3 compares the principal results of this and prior data. The performance results at the top present the variances of the task error and control force, along with their partioning into: target, disturbance and remnant components. Remnant is the uncorrelated, or noise, portion of each signal. For convenience, an RMS scale is also noted on the right hand of each plot.

In both experiments, the magnitude and the partitions of performance measures were practically identical. The control-force remnant of roughly 40 percent is reduced to 15 to 20 percent in the error signal, because the controlled element filters the high frequency remnant components.

The primary measures of pilot behavior are the "opened loop" describing functions for the simultaneously-closed-loops involved in target following and in disturbance suppression, shown at the bottom of Fig. 3.* The points are from the

*These "opened-loop describing functions are derived in Ref. 1 which also discusses their relevance. Suffice it to say, here, that the "opened-loop" transfer functions have the same significance in this multi-loop situation as the conventional "open-loop" transfer functions do for single-loop situations. Roughly K/s-like behavior is expected if the crossover law is followed. The stability margins and bandwidths can be read from such plots in the conventional manner.

a) PERFORMANCE



b) BEHAVIOR

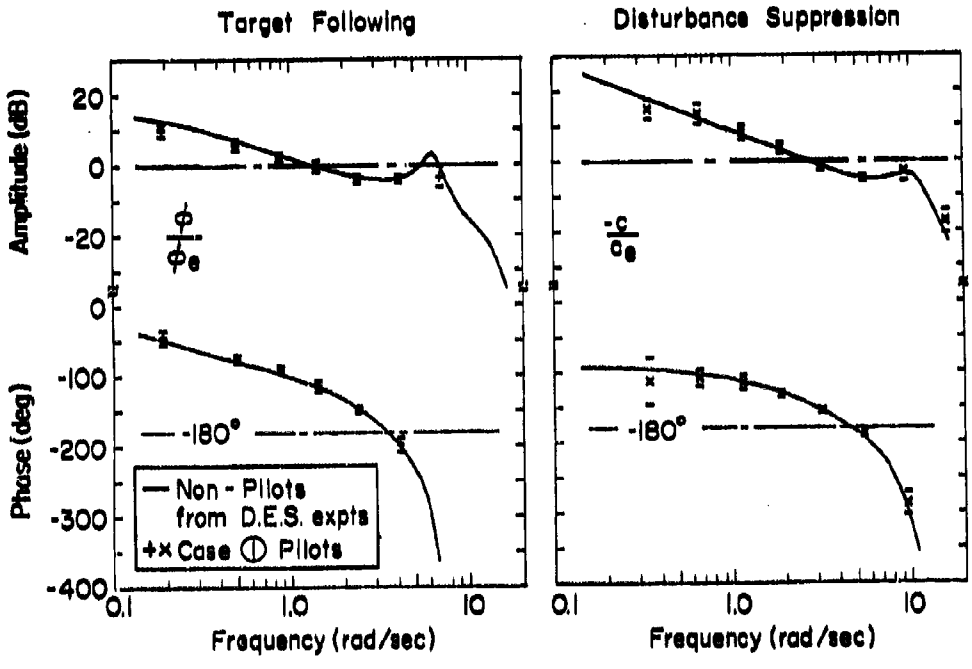


Figure 3. Results for Tie-In with Prior D.E.S. Experiments

present LAMARS experiments, as shown by the mean and standard deviations for the 3 pilots (at least 2 runs each). The low variability shows that all pilots followed the same behavior. The curve through the points is not a fit to this LAMARS data but is, in fact, the fit from the previous DES experiments with non-pilots. One could hardly expect a better fitting curve for the present data, and an independent fit (not shown) gave nearly identical parameters. It is apparent that the three experienced pilots in the present LAMAR simulation adopted identical behavior and performance as that of the non-pilots in the previous DES experiments.

We further conclude that the behavior seen in the previous experiments was, in fact, driven by the combination of controlled element, input and motion properties, because it is universally adopted by both well practiced non-pilots and by skilled pilots as well. This has two important implications:

- Experiments performed on the LAMARS may be tied in closely with those performed on the DES, thereby improving the breadth of coverage in basic research problems such as this.

- The DES could be used to train pilots or to perform preliminary experiments for the LAMARS, thereby reducing the total cost, because the DES is cheaper to operate than the LAMARS and may be more available.

The universality of pilot adaptation and performance demonstrated here gives further impetus to the proposition in Ref. 1, that valid control-theory models of multiloop pilot behavior should closely match these results. All the inputs, controlled elements and washouts are representable by simple pole-zero transfer functions, as can the pilot's behavior (see Refs. 1, 2 for details). The challenge is this: can the optimal or classical models of pilot behavior replicate these results with a consistent set of cost functions and adjustment rules?

Effects of Sway Motion

Figure 4 shows the effects of freeing the sway degree-of-freedom from roll-only (Case (9) vs. Case (7) ; both for the "smoothed" controlled-element). At the top left, the sensed lateral-specific-force is shown. In the roll-only case (which corresponds to infinite washout) the maximum spurious LSF reached approximately 0.1 gy, with peaks of 0.2 to 0.3 g's. With lateral sway freed and medium washout, these peaks were reduced by about 20 or 30 percent. The achieved roll rate was practically identical, showing that the pilot used roughly the same high frequency control actions. The sway travel was zero in the roll-only case and just under 4 ft in the medium case (occasional peaks to 10 ft). The tracking error performance was practically identical in both cases.

The bottom of Fig. 4 shows the effects of sway on the pilot control behavior. The dashed curve is that fitted to the roll-only data (not presented) for the smoothed control element. The solid-curve is that fitted to the given data points using the Model Fitting Program described in Ref. 1. The refined pilot model included his perception of both roll rate and roll accelerations, and LSF cues in the motion path, in addition to the classical pilot lead and time delays of the visual path and second-order neuromuscular dynamics. As in all of the DES

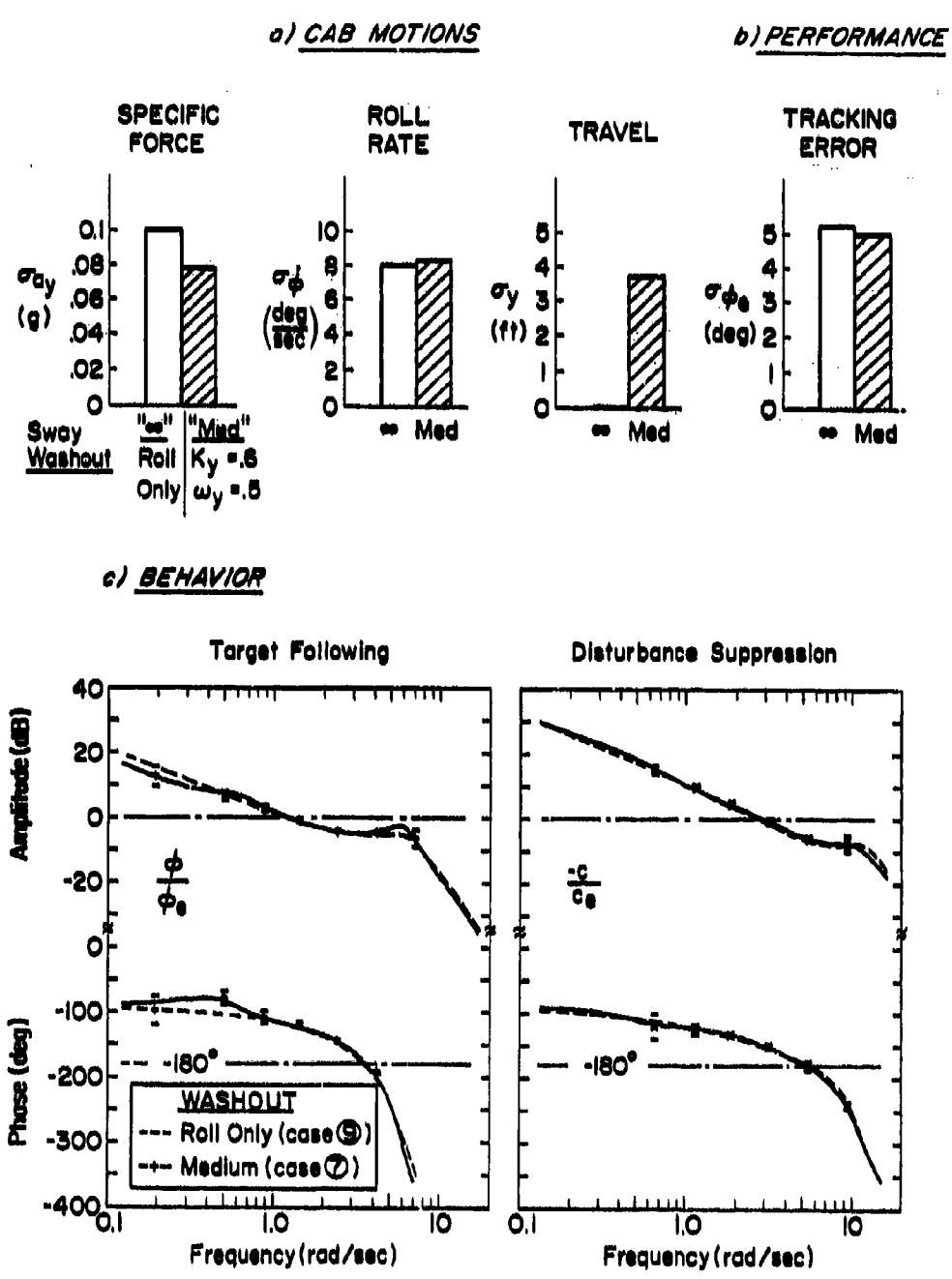


Figure 4. Effects of Washed-Out Sway vs. Roll-Only

and LAMARS cases analyzed to date, the curve fit to the data points is excellent, and the model clearly captures all the nuances of the measured describing functions.

Figure 4c shows two main effects of washed-out sway motion:

- The disturbance-control loop ($-c/c_e$) is essentially unaffected because neither the motion cues (which dominate $-c/c_e$ at high frequencies) nor the visual cues (which dominate at low frequencies) are affected by the medium washout.

- The target tracking loop (φ/φ_e) is only affected at low frequencies, where the lateral-specific force cues give helpful tilt cues in the roll-only case, but are distorted and less effective in the washout case (as shown by the lower gain below $\omega = 0.5$ rad/sec).

On the other hand the target following transfer function shows a little lower gain at the very lowest frequencies implying less tilt cue which has been shown previously in Ref. 1 to be a favorable effect.

Overall, there is remarkably little difference in performance and behavior between the roll-only and the roll-plus-sway case, with this medium sway washout. One reason is that the washout did not remove much of the tilt cue, as shown by the similarity of LSF in Fig. 4a. The pilot's subjective comments are discussed separately, later.

Effects of Various Degrees of Sway Washout

The main set of variables, in which ω_y and K_y were varied produced almost negligible changes in the specific measures of pilot performance and behavior. It had been expected that the wide range of sway washouts (which do result in adverse comments) would result in a behavioral change, but this was not apparent from the data. In general, the describing function differences among the sway washout combination (Cases ③, ⑥, ⑤, and ⑧B.4) were similar to the data separations shown in the previous Fig. 4. The one case with significant reductions in describing function gains, ②, was highly confounded by encounters with soft stops (travel limiters) hence no transfer function data will be presented.

Figure 5 compares various measures of motion output and performance among the various sway washouts tested, arranged in the same order of increasing severity as in Table 1. The first set of bars compares the lateral travel for the various washouts. As would be expected, the RMS travel reduces, from the untested (but calculated value) of over 40 ft for free flight, to 3.3 ft (10 ft peaks) for the least sway washout that could be used, to further reductions to ≈1 ft (3-4 ft peaks) for the medium and high degrees of washout. The extent of travel is greatly reduced by the sway washout, as was expected.

The second set of bars shows that the spurious lateral specific force cue increases with washout severity as was expected. The RMS levels of this LSF are quite low, with the peaks reaching values of only 0.1 g for the medium, high, and roll-only cases. The basic tradeoff between travel and miscue is clearly apparent from these data: one obtains reduced travel at the expense of increased miscues.

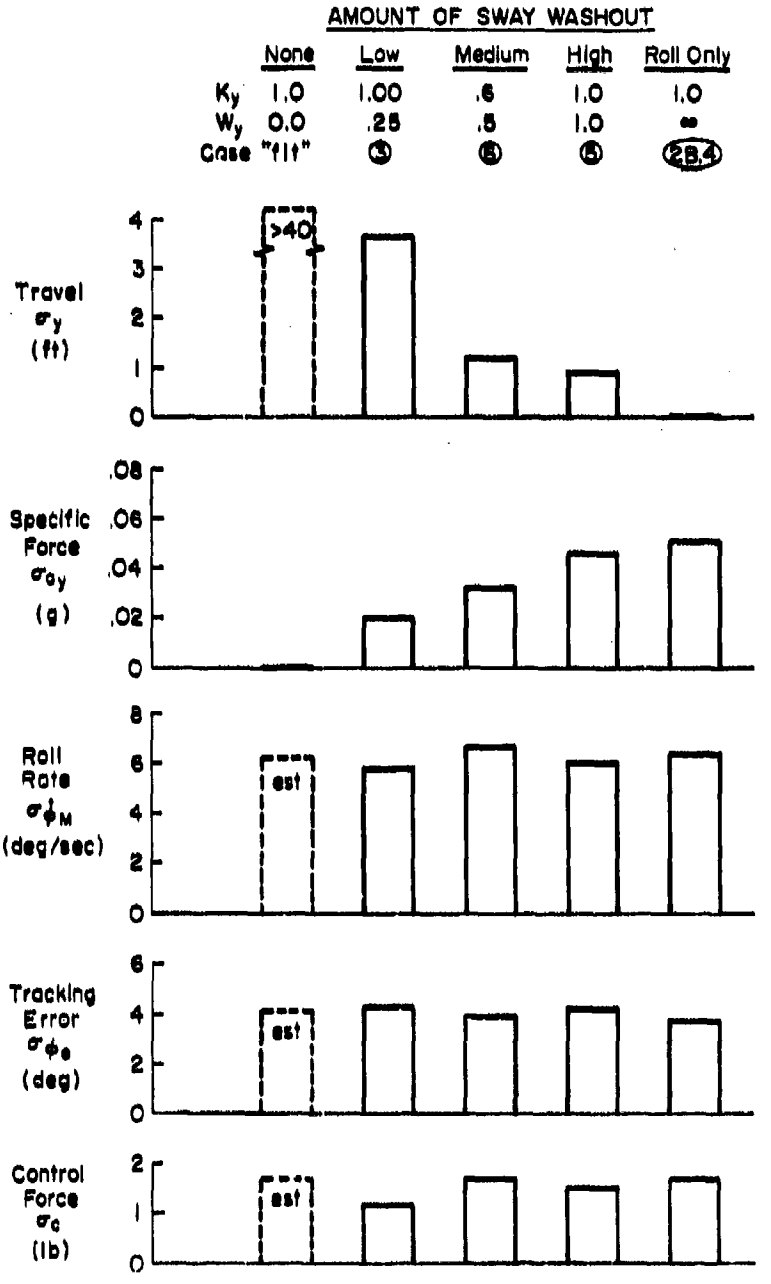


Figure 5. Comparisons Among Various Sway Washouts

Looking now at the bottom of Fig. 5 for measures of the tracking performance and behavior, one sees that the roll rate was practically identical among all sway washouts. This means that the roll loop was closed in the same manner, regardless of the presence of the (often sub-threshold) LSF cues. In other words, lateral specific force cues were essentially ignored in the roll loop closure.

The same conclusion applies to the roughly equal tracking errors and control force measures at the bottom of Fig. 5. These performance measures cannot be considered to be the whole story because the pilot's subjective comments differed significantly among these various washouts, as will be discussed later.

A word of caution is in order, here as might be implied that if reliable behavior can be obtained with roll-only simulation considerable expense (in terms of lateral motion systems) could be saved by eliminating the sway motions. However, lateral specific forces are often simulated for other purposes for roll-yaw coordination (e.g., for rudder use, or engine-out effects, or direct-side-force simulation) so roll-only simulation is not a panacea for the general case.

Another problem was hinted at during these tests: that of possible negative transfer of training from the roll-only case (wherein a rightward-tilt requires a left aileron correction) to the free-flight case (where the corresponding leftward LSF in general would not result in a left aileron correction). A hint of this was present when the sway-degree-of-freedom was turned on, so the roll-only runs were discontinued as warmups for the sway washout sessions. This bears further study.

Linear vs. Non-Linear Washout

A nonlinear washout algorithm was evolved which would automatically adjust the sway washout filter frequency, as a function of the sway command states, as well as the proximity to the sway limits. Soft limiting of large excursions can only be obtained by the addition of LSF miscues and it was the basic question of this comparison to see whether the decelerations for a large motion would be acceptable to the pilot. This algorithm keeps the washout as low as possible during low roll motions, and increases it only as much as is necessary to cope with large roll motions. Consequently, one algorithm could be used for a large number of conditions, instead of having to "optimize" the sway washout parameters for each condition, as is necessary with fixed sway washouts.

Figure 6 depicts the results of comparing the nonlinear with the "medium" washout, which it approaches for small motions. The larger reference input was used, to maximize the sway cues. The bar charts for cab motions show that the LSF, roll rate, and lateral travel were practically identical for either case. What is not shown here is that the non-linear washout case (shown shaded) resulted in fewer bumpings of the parabolic limiters than did the linear case. When the cab encounters the parabolic limiters, large braking forces are put in, and the yaw rate coordination mechanism of the cab is ineffective, so that there is a tendency to become disoriented after each encounter. These incidents were greatly reduced in the non-linear washout case, which the pilots liked. On the other hand, the adaptive circuitry puts in slightly higher smooth decelerations (which are miscues).

The tracking error plot shows that the performance was practically identical in either case. Figure 6c gives the transfer functions comparing the fitted curve for the linear case (Case (7)), with the non-linear data of Case (8c) showing practically identical behavior. (A linear model fit for the non-linear data is not strictly

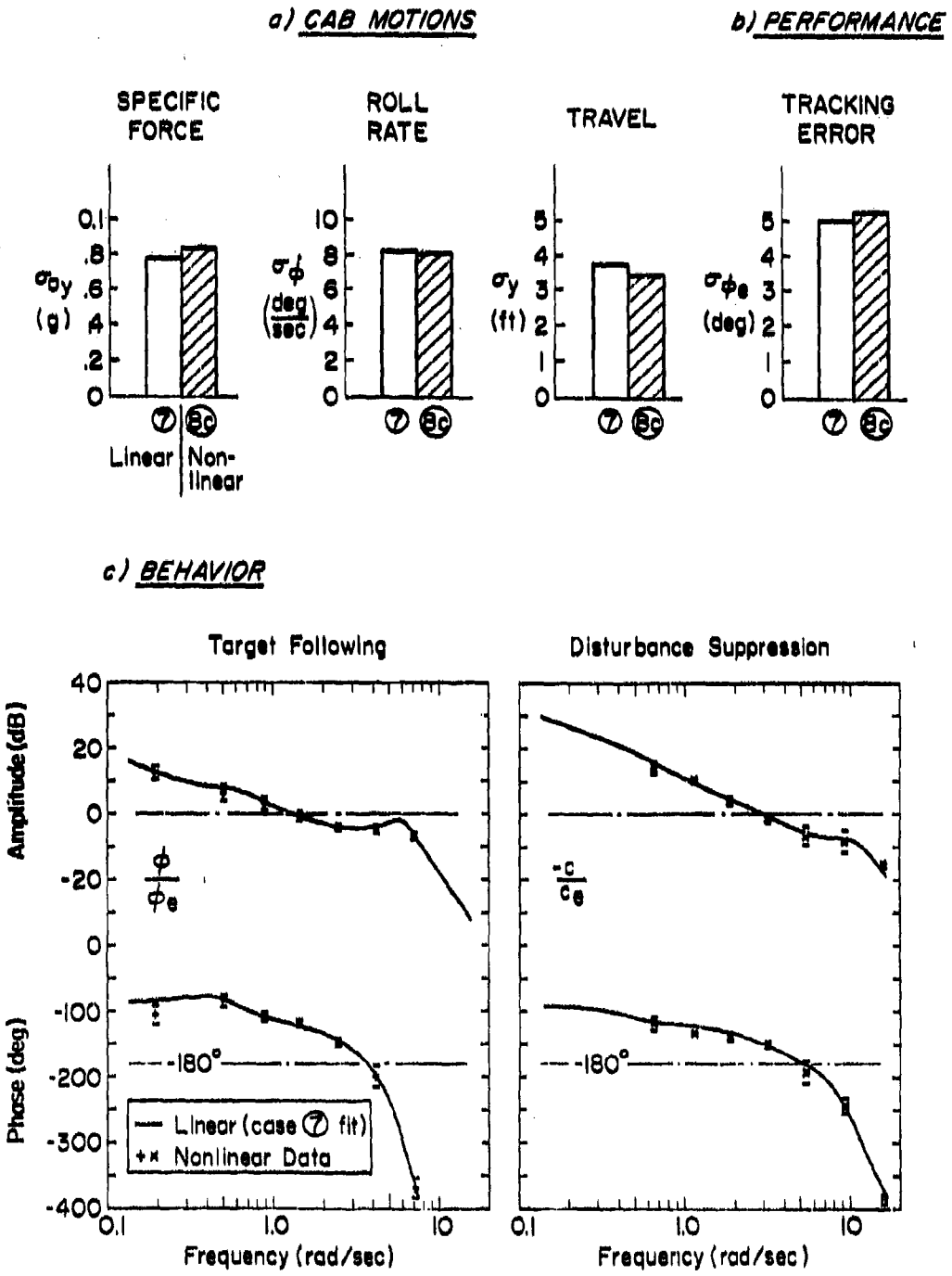


Figure 6. Comparison of Linear and Nonlinear Washouts

applicable because of the time variation in the washout properties). Nevertheless, it can seem that the linear-case curve fits the nonlinear-case data quite well, implying essentially identical behavior.

The implication of these results is that the nonlinear sway washout is a promising way to economize on experimental operations because the algorithm seems to work over a large range of inputs. A number of discrete (i.e., bank and return) and other input cases were checked with the nonlinear washout during this experiment; and, in all cases, the results seem to be similar to those shown. We recommend this nonlinear sway washout scheme for other simulators having similar travel limits.

CORRELATION OF SUBJECTIVE AND OBJECTIVE DATA

Although a great deal of pilot commentary was taken during these experiments, it was difficult to establish firm correlations with the washout parameters because the motions, although noticeable, did not strongly affect the pilot's control strategy and, in many cases the LSF involved was small enough to be almost "subthreshold." It turned out that this LSF threshold effect was the dominant distinguishing feature between cases which could be consistently evaluated and in cases which were vaguely and inconsistently evaluated. As noted earlier in Fig. 5b the rms LSF were under 0.05 g for the reduced input cases, and approached 0.10 for the reference input cases. The peak LSF did vary somewhat more than this due to occasional large motions; this enabled the pilot commentaries to be made when they occurred. In many of these cases the pilot was also allowed to make bank-and-return maneuvers within the confines of the simulator; in such cases the lateral specific forces often exceeded 0.1 g's so the sway washout effects became more readily apparent.

Accordingly, we have plotted the peak LSF vs. the washout filter frequency (ω_y) and washout attenuation factor (K_y) on Figures 7a and 7b, respectively. Along with each of the data points is given the consensus of comments among the pilots. First, notice the black points which are those obtained during the roll tracking cases with reduced inputs. The peaks seldom exceeded 0.1 g's and the general consensus of pilots comments was vague. Only for very low values of K_y in Fig. 7b (which correspond to roll alone) did the LSF exceed 0.1 g's, and then was some comment about noticing the tilt cue or the "leans."

Considering next the open symbols for the rapid bank and return maneuvers, it can be seen that, when the lateral specific force peaks exceeded about 0.1 g's, there was distinct comments which had some correlation with the task variables. (Cross plots of these effects will be shown later).

It is inferred from careful study of these comments that: when lateral specific force peaks lie about below 0.1 g's, the LSF effects are, at best, only vaguely perceived and, at worst, are inconsistent. This hypothesis may help explain some of the apparently inconsistent results from other investigators, where the lateral specific force peaks are known to have been less than 0.1 g (e.g., Refs. 5 and 6). Others (e.g., Refs. 7 and 8) have suggested that such an "indifference threshold" would be operational on lateral specific force cues, and our findings verify these suppositions.

The findings have important consequences for the design and interpretation of experiments involving lateral motions. If the expected LSF cues are less than

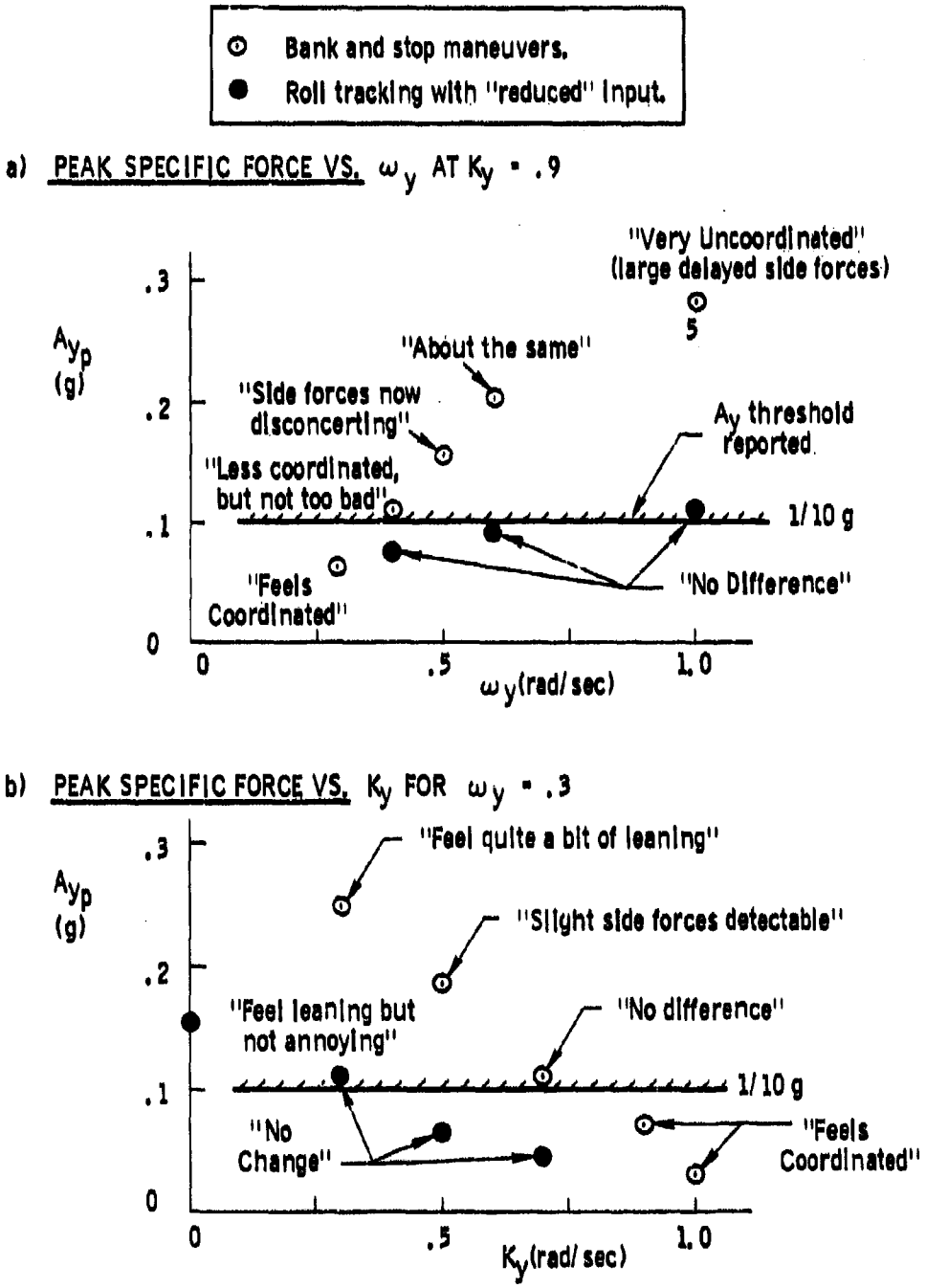


Figure 7. Peak Specific Force and Comments vs. Sway Washout Frequency and Gain

0.1 g's, the experiment may be expected to encounter confused pilot commentary. If the small cues are used in important ways for coordination and for detection for certain types of failure, then such small levels might be worth investigating. On the other hand, if the lateral specific force cues are greater than 0.1 g's they will surely be noticed and used, and if they are due to motion-base artifacts, they may affect the results in a negative manner.

Gathering and sifting those cases where the motion cues were strong enough to give significant commentary, the correlations shown in Fig. 8 summarize the consensus from this experiment. Figure 8 represents a "subjective commentary map," wherein the coordinates are washout frequency along the abscissa and attenuation factor as the ordinate. The bottom left corner represents fully-coordinated free flight. Regions of distinct subjective effects are denoted by the fuzzy boundaries and paraphased comments.

It is noted that a good portion of the interesting washout range (values of K_y near 1.0 and $\omega_y < 0.2$ rad/sec) exceeds the travel limits of the simulator. A barely acceptable region was observed for $K_y = 0.5 - 0.7$ and $\omega_y = 0.2 - 0.4$ rad/sec, where the LSF miscues are small enough and the peaks are not unduly delayed with respect to the roll angles to cause apparent distortion.

Consider next the "leans" region (at the top) where the attenuation is reduced towards zero, implying no sway motion whatsoever. In these cases, the motions approach roll-only and the tilt cue is closely in phase with the roll angle. Hence, there is little confusion between the source of the lateral specific force and the observed roll angles. We call this syndrome of effects "the leans" because the pilots are aware that the perceived LSF is due to a leaning effect of the simulator.

At the bottom right of Fig. 8, the attenuation factor is nearly unity, and high washout frequencies are present, the "delayed side-force" effect (noted at the start on Fig. 1) was apparent. Here, the side force peak, even though small, was delayed with respect to the roll angle peak. This seemed to cause some disconcerting motion effects, because it is harder to assimilate into the pilot's experience with aircraft. Finally, at the right center there is a large range of undesirable properties where the washout filter frequencies are on the order of 1.0 rad/sec and attenuation was somewhat reduced (K_y 0.5 - 0.9) which produced highly distorted lateral motion cues. In many such cases the pilots claimed that the spurious cues were like an inexperienced student pilot putting improper inputs into the rudder pedals. The obvious conclusion from this Fig. 8 is that: it is difficult to achieve an acceptable degree of washout within the ±10 ft confines of the LAMARS Simulator. This conclusion is not inconsistent with experience of other aircraft simulations both on the LAMARS and elsewhere. So much landscape is required to reduce the spurious lateral specific force cues below 0.1 g's that acceptable sway simulation can seldom be afforded in a practical simulator.

CONCLUSIONS

The following conclusions are drawn from these experiments.

1. There is an excellent tie-in between the dynamic environmental simulator and LAMARS simulations when the same case is simulated.

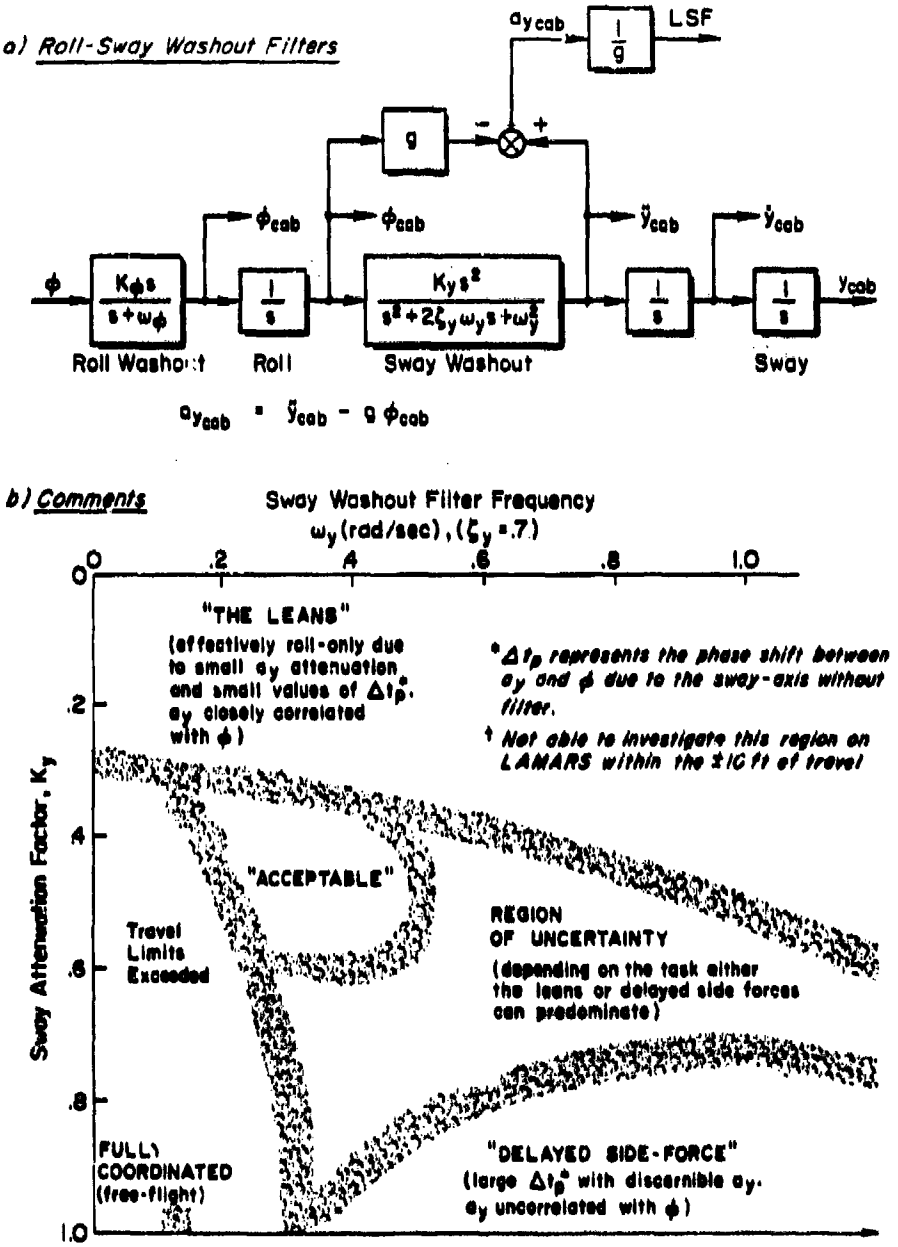


Figure 8. Correlation of Comments vs. ω_y and K_y

2. The present pilots and trained non-pilots from earlier experiments showed nearly identical behavior and performance, implying universality of adaptation and results. However, the pilots needed much less training time than the non-pilots due to their previous flight experience. This points up one advantage of using actual pilots in such basic research; the amount of training time necessary to obtain valid results can be reduced by an order of magnitude when real pilots are used, thereby offsetting their scarcity and expense.

3. Both step and random tracking gave rise to spurious lateral-motion-cues (the turn-coordinated free-flight case would have none) which were fuzzily categorized as "out-of-phase," "like a student on the rudder pedals," etc. Analysis showed these to be roughly correlated by time-and-frequency-response parameters related to K_y and ω_y.

4. Neither the roll tracking behavior nor error performance were significantly affected by a variety of lateral-sway washouts.

5. Pilot comments were consistent only when the peak lateral specific forces exceeded about 0.1 g. Pilots noticed "leans" (tilt effects and "delayed side forces" (delays between peak roll angles and peak side-forces).

6. A nonlinear beam washout filter (working on computed sway states) reduced the rate of soft-stop encounters, at the expense of occasional, smooth lateral-specific-force (a_y) peaks, but otherwise did not affect behavior or performance. It promises to provide an adaptive washout which does not need to be "fine-tuned" to avoid hitting stops while minimizing supurious washout artifacts. Additionally, it should be especially useful during training, where motion cue usage is changing.

7. These results imply that sway motions well over ±10 ft should be provided to reduce lateral motion miscues to acceptable levels for realistic discrete or random tracking inputs.

   It would be interesting to find and simulate exactly some real-world tasks in which the free-flight motions would be under ±10 ft, such as aerial-refueling or shipboard VTOL landing. Our hypothesis would predict that, unless the lateral specific forces would exceed 0.1 g (unlikely), lateral motion cues would not be significant in such cases.

REFERENCES

1. Jex, Henry R., Raymond E. Magdaleno, and Andrew M. Junker, "Roll Tracking Effects of G-Vector Tilt and Various Types of Motion Washout," STI Paper No. 226, presented at Fourteenth Ann. Conf. on Man. Control, NASA CP-2060, Nov. 1978, pp. 463-502.

2. Hass, R. L., H. E. Hotz, and G. R. Mills, "The Large Amplitude Multi-Mode Aerospace Research (LAMAR) Simulator," AIAA Paper No. 73-922, presented at AIAA Visual and Motion Simulation Conference, Palo Alto, CA, Sept. 1973.

3. Jewell, Wayne F. and Henry R. Jex, A Second Order Washout Filter With A Time-Varying Break-Frequency, STI WP 1094-9, Revised Mar. 1979.

4. Levison, W. H., S. Baron, and A. M. Junker, Modeling the Effects of Environmental Factors on Human Control and Information Processing, AMRL-TR-76-74, Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio, Aug. 1976.

5. Shirley, R. S., "Motion Cues in Man-Vehicle Control," M.I.T., Cambridge, MA, Sc.D. Thesis, Jan. 1968.

6. Stapleford, Robert L., Richard A. Peters, and Fred R. Alex, Experiments and a Model for Pilot Dynamics with Visual and Motion Inputs, NASA CR-1325, May 1969.

7. Sinacori, John B., A Brief Survey of Motion Simulators' Drive Logic with with Emphasis on the Roll Axis, STI SP 1094-2, May 1977.

8. Hofmann, L. G., and Susan A. Riedel, Manned Engineering Flight Simulation Validation. Part I: Simulation Requirements and Simulator Motion System Performance, AFFDL-TR-78-192, Feb. 1979.

ACKNOWLEDGEMENTS

This work was sponsored by the USAF Aerospace Medical Research Labs, Environmental Medicine Division, at Wright Patterson AFB, Ohio under Contract F33615-77-C-0508. Andrew Junker was the Contract Technical Monitor and a Principal Investigator. Appreciation is extended to: John Sinacori (then at STI), for assisting with the washout/LAMARS drive logic interface; Marvin Roark of AMRL, for adapting the AMRL measurement and pilot describing function programs to AFFDL computers; John Bankovskis of AFFDL, for handling these programs at AFFDL and for modifying the LAMARS drive logic and operating mechanizations to permit on-line variation of the washout parameters; Jim Ater and Warren Miller of AMRL who performed most of the training and test operations on the LAMARS, and to the four anonymous military pilots at AFFDL, who participated in these experiments, often on their own time.

CONTROL LOADING METHODOLOGY USING A MICROCOMPUTER

Dr. Gerry Albers
Aerospace Engineering
University of Dayton
Dayton, Ohio 45469

ABSTRACT

The tactile frequency response of a human is much higher than the visual response. A specific example of this characteristic can be witnessed in the simulation of an aircraft. The visual scene content is perfectly acceptable when the image is refreshed at a rate of 30 times per second. However, the pilot will feel a definite vibration or "hum" in his control stick (particularly while the stick is being moved) if the forces on that stick are computed at the same rate. Furthermore, whenever a sharply rising force is encountered such as at a travel limit or at a detent, high frequency components must be present or else the "sharpness" is lost. Therefore it is essential in a simulator that an independent computer be provided to compute the forces which the pilot feels, since the host computer is usually "framing" 15 to 30 times per second.

The various components of the control loading forces are discussed in this paper. A brief history of control loading implementations is also presented with the emphasis on two modern computational approaches, one of them analog, the other digital. The digital approach employs a microcomputer system which is discussed in detail. The hardware sub-systems were off-the-shelf components and the software was written in a high-order language. The equations and their implementations are discussed, with particular attention paid to the frequency compensation necessary to produce "sharpness". A high iteration rate was achieved and the system was capable of solving the force equations for all three axes (pitch, roll and yaw) in the allotted time. Recommendations for future implementations are also presented. This digital control loading base was the result of a research effort with the Aeronautical Systems Division at Wright Patterson Air Force Base, Ohio.

AN APPROACH TO DESIGN OF SIMULATION SYSTEMS FOR TRAINING AND SELECTION OF PILOTS IN AN ENVIRONMENT OF COMPETITION

Leonid Lipchin, Ph.D.
Higher Order Software, Inc.
Cambridge, Massachusetts 02139

Abstract

This paper describes training and selection methodology for developing maximum efficiency of an operator in the context of compensatory tracking. The major idea is a motivation principle of training based on competition. The function of operator efficiency (OEF) is introduced. The classification criterion for dynamic stereotypes of operators is discussed. A particular competitive structure is chosen for each dynamic stereotype. While competing for the leadership position, (operator selection requirements) the operator continually compares the relative dynamics of his OEF to others.

The test runs on the condition that the operator believes he competes against authentic recorded time history OEF's of real operators who have been previously tested. In reality, his display is indicating computer-simulated data which can be manipulated by the examiner depending on the operator's performance. Application of this methodology to training and selection of cosmonauts for the docking task is considered.

Introduction

Proper training and selection of operators in manual control is one important method to make a man-machine system more efficient.[1] This problem has been studied in the context of compensatory tracking where the operator must determine his control actions from indirect information. For specific control tasks, the operator's dynamic stereotype develops during the process of training. The dynamic stereotype describes the operator's skill to choose and maintain the compensatory control function during a period of time. We can observe the ability (qualitative parameters) of people to control a system through the parameters of skill (quantitative parameters). Physical skill and psychological factors, important in man-machine interaction, form the dynamic stereotype as discussed by others.[1]

We can construct a mathematical function E(t) of operator efficiency (OEF) for each specific task. This function takes into account how well the subject succeeded (or failed) throughout the task step by step. The final OEF value E(t) is a measure of success integrated over the entire task.

The approach is based on the following statements.

(1) For each specific task, operators can be classified by a finite number of dynamic stereotypes

$$D_i,\ 1 \le i \le K. \qquad (1)$$

(2) For each dynamic stereotype, a definite range of possible maximum value of the final OEF can be determined:

$$E_{min} \le E(D_i,T) \le E_{max}. \qquad (2)$$

(3) The operator expresses his dynamic stereotype through the time history (tempo) and final value of the OEF, $E_{op}(D_i,T)$.

(4) The final value of the measured OEF for any subject can approach the maximum of the range for his stereotype only through proper training.

(5) Motivation based on competition is the most important factor in training efficiency.

(6) There are three ranking stereotypes, L, which determine the best initial ranking of any subject relative to his competitors: (a) LEADER -- one who achieves maximum efficiency only when he starts by leading his competitors; (b) OUTSIDER -- one who achieves maximum efficiency only if he begins behind all other competitors; and (c) NEUTRAL -- one who achieves maximum efficiency irrespective of initial ranking.

If these statements are accepted, it is possible to establish training and selection methodology for developing maximum efficiency of any operator.

Methodology

Basic Idea

While competing for leadership position (operator selection requirements) the operator continually compares the relative dynamics of his OEF to others. For this methodology display of competition, which can indicate the OEF time history of all competitors to subject and examiner is required. This can be added to existing cockpit displays. The test runs on the condition that the operator believes he competes against authentic recorded time history OEF's of real operators who have been previously tested. In reality, his display is indicating computer-simulated data, which can be manipulated by the examiner.

First (Fig. 1), it is necessary to collect and process training test data for the specified task from a large number of operators over a period of time. From a model of the task, the OEF and the required final value for success in the task are determined. Through analysis of this data, the number of dynamic stereotypes as well as the characteristic time history and range of final OEF values of the dynamic stereotypes can be established (Fig. 2). A few average time histories with different characteristic frequencies for each dynamic stereotype can then be modeled by computer methods.

Best Available Copy

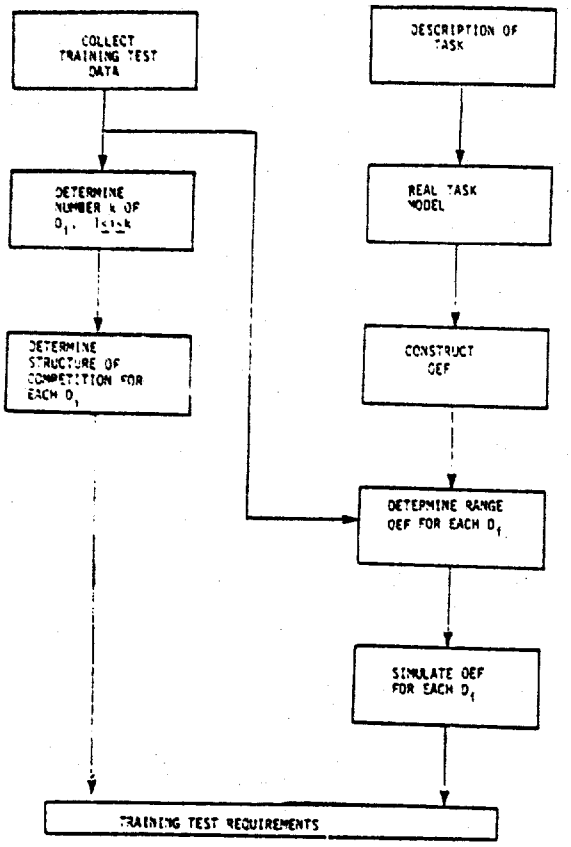
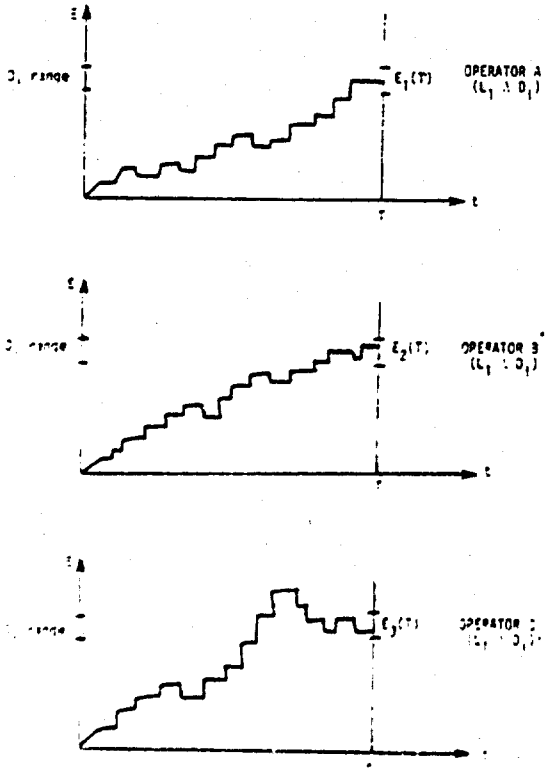


Fig. 1 Flow Chart of Major Tasks Performed in Developing Training Test Requirements.

$E(L_j \wedge D_i, t)$ - CURRENT OEF VALUES

$E(L_j \wedge D_i, T)$ - FINAL OEF VALUES



The examiner can create competitors of any dynamic stereotype by computer simulation, i.e., he can compose any environment of competition for any operator. The examiner must first (Fig. 3) determine the subject's ranking stereotype by a psychological test. The subject is instructed that selection will be made on the result of competition. He is told that some of his competitors have passed the examination before. He is instructed about the meaning of the OEF and how it is displayed, and he is told that he can continuously see the time histories of OEF's of his competitors which have been previously recorded. He is instructed that he must place first in the competition to be selected. Then the training begins.

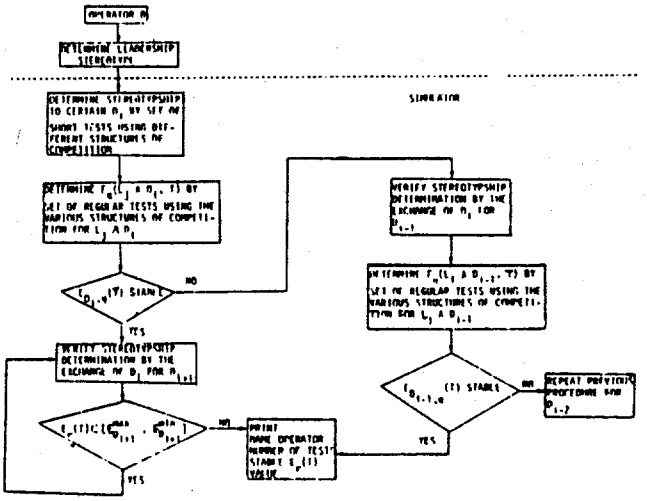


Fig. 3 Flow Chart of the Training Processing Algorithm.

First, the examiner runs short tests to determine dynamic stereotype based on leadership stereotype ($L_j \wedge D_i$). During this time, the examiner simulates different structures of competition and analyses the time history of OEF of the operator. Then, the normal test is run where the operator is compared to some number of simulated candidates. The test is finished and the examiner analyzes test data. The next test will be for a more complicated structure of competition. Stability of final OEF value $E_{jj}(L_j \wedge D_i, T)$ during a few tests can reveal: (1) maximum efficiency of operator; and (2) if a mistake was made in initial dynamic stereotype classification. Further tests can then be made to confirm new classification. We can say that the operator is sufficiently trained if the final OEF value has been stabilized and also has reached its threshold value over the number of running tests.

Application

This methodology was invented[3,8,9] and patented[2] under sponsorship of Project "SOYUZ's" in the USSR. It was applied to the training and the selection of cosmonauts for the docking task.[3,4] Figure 4 shows the geometry of the guidance plane. Relative velocity

Best Available Copy

$\dot{\rho}$ characterizes the time remaining until docking and lateral velocity, V_{lat} determines the projected error h. The task of the cosmonauts is to: (1) reduce $\dot{\rho}$ to a minimum; (2) eliminate initial error; and (3) keep center axes of both spacecrafts the same.

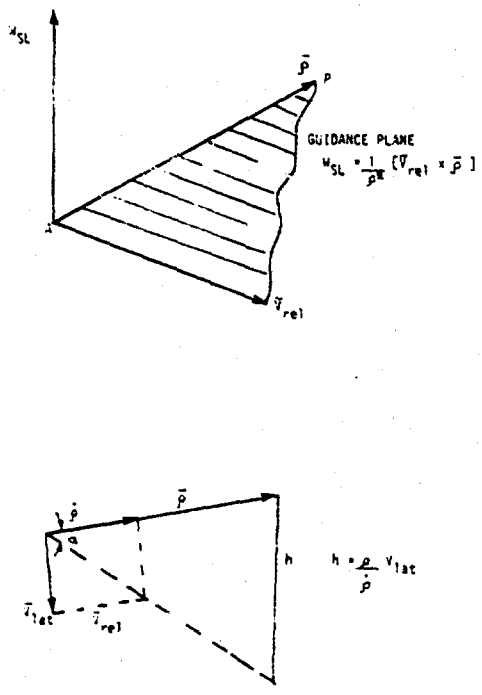


Fig. 4 Geometrical Structure of the Docking Task.

Figure 5 shows the initial and final conditions for docking. For visual determination of mutual orientation of both vehicles on the passive spacecraft there are landmarks and spotlights. By these aids, the cosmonaut can determine declination between mutual central axes in one plane, as well as, "top" and "bottom" of coupling.

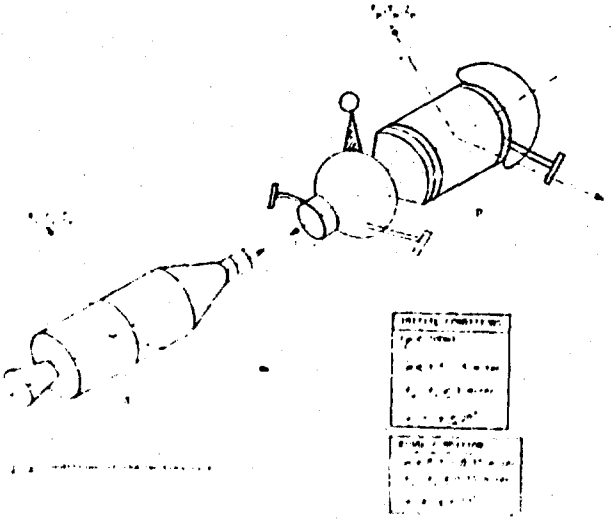


Figure 6 shows the structure of control of spacecraft during the manual docking manuever. The cosmonaut makes decisions about required control actions only on the basis of information about distance to target and relative velocity from a radar system, and bearing and roll from an optical sight system. The cosmonaut manages a set of rocket engines in order to control approach parameters as well as vehicle stability. Because the cosmonaut has a finite response time, he brings into the control system a minimum amount of angular instability, ω_y, ω_z.

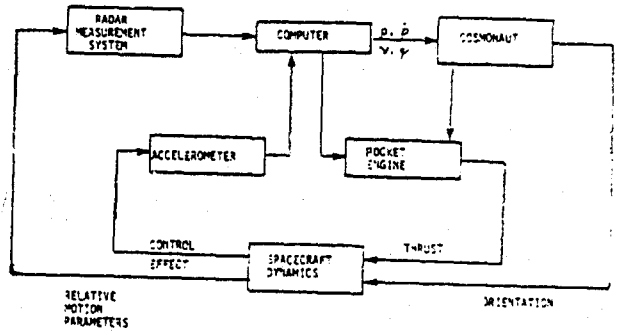
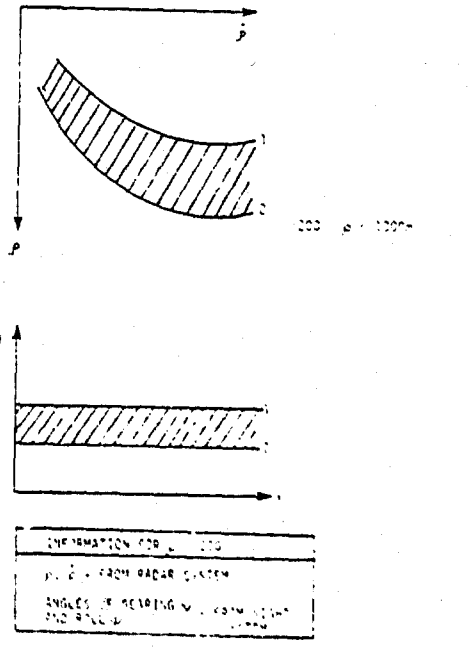


Fig. 6 Control Model of the Manned-Spacecraft System.

For the first part of the docking manuever (ρ = 200 ÷ 1000 m) the cosmonaut should keep the parameters of approach within the proper hatched field which is displayed in Fig. 7. The limit of the field is modeled by computer, but during the last part of docking (most important, ρ < 200 m), the cosmonaut has more responsibility to minimize the error in approach parameters. The OEF for this task was determined on the basis of a mathematical model of docking, which determines the time and dynamics of docking and expenditure of fuel F_x and F_y.



Best Available Cop

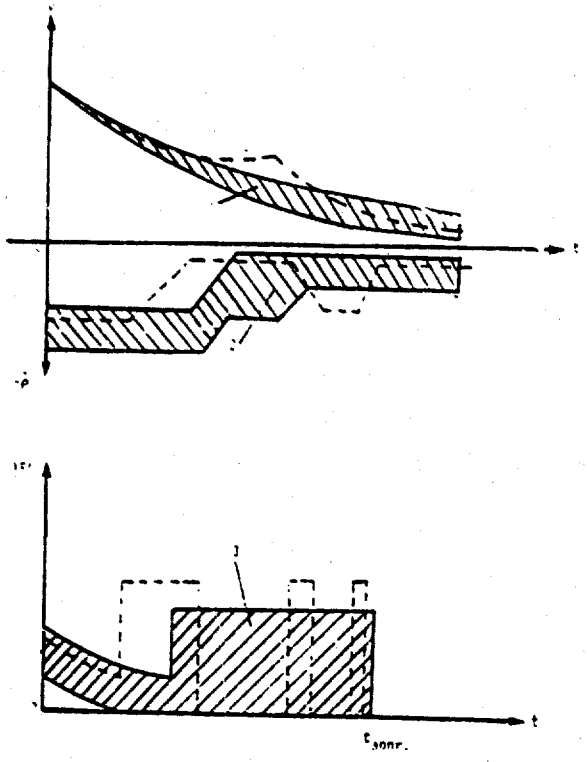


Fig. 8 Permissible Fields of Control for Final Conditions of the Docking Task

Figure 8 shows permissible ranges of control (hatched areas) for distance (1), velocity of approach (2), and fuel consumption rate (3) as functions of time. Typical control trajectories for operators during simulation are shown by the dashed line. Excursions outside of the permissible control fields (PCF) characterize non-optimal control behavior by the human.

The Bush-Mosteller function was chosen to construct the OEF:[11]

$$E_i = \alpha E_{i-1} - (1-\alpha)\, x_i \tag{3}$$

$$x_i = \begin{cases} 1, & F \subset PCF \\ 0, & F \not\subset PCF \\ E_{i-1}, & F_i = F_{i-1} \end{cases} \tag{4}$$

For this specific task the function has the following form:

$$E(T) = \frac{1}{T_o} \sum_{i=0}^{N} \exp\{-\frac{T-\Delta t_i}{T_o}\}\, x(\Delta t_i)\, \Delta t_i \tag{5}$$

where[9] $T_o = \Delta t/(1-\alpha)$ and $\alpha = [1 - (4\Delta t/T)]$.

This function was constructed using a Kalman-Bucy filter. The author developed a digital scalar form of this filter which, for this specific purpose, did not require matrix inversion.[10] Assume that the components of the n-dimensional vector $\vec{E}$ are phase coordinates of the scalar function E(t), which vary with time. Then the discrete variant of Kalman-Bucy estimator:

$$E_n = \Phi E_{n-1} + \Gamma \omega_n \tag{6}$$

$$Z_n = H E_n + \nu_n \tag{7}$$

and its recurrent equations,

$$\hat{E}_n = \Phi \hat{E}_{n-1} + P_n^- H'(H P_n^- H' + R)^{-1} (Z_n - H\Phi \hat{E}_{n-1}) \tag{8}$$

$$P_n^+ = P_n^- - P_n^- H'(H P_n^- H' + R)^{-1} H P_n^- \tag{9}$$

$$P_n^- = \Phi P_{n-1}^+ \Phi' + \Gamma Q \Gamma' \tag{10}$$

will be as follows. Instead of (8), we will have a system of scalar equations:

$$\left.\begin{aligned} \hat{E}_1[n] &= E_1^-[n] + \frac{P_{11}^-[n]}{P_{11}^-[n] + \sigma^2[n]} (Z_1[n] - E_1^-[n]) \\ \hat{E}_2[n] &= E_2^-[n] + \frac{P_{12}^-[n]}{P_{11}^-[n] + \sigma^2[n]} (Z_1[n] - E_1^-[n]) \\ &\cdots\cdots\cdots\cdots \\ \hat{E}_m[n] &= E_m^-[n] + \frac{P_{1m}^-[n]}{P_{11}^-[n] + \sigma^2[n]} (Z_1[n] - E_1^-[n]) \end{aligned}\right\} \tag{11}$$

Similarly, instead of (9) we will have:

$$\left.\begin{aligned} P_{11}^+[n] &= P_{11}^-[n] - \frac{P_{11}^-[n]}{P_{11}^-[n] + \sigma^2[n]} P_{11}^-[n] \\ P_{12}^+[n] &= P_{12}^-[n] - \frac{P_{11}^-[n]}{P_{11}^-[n] + \sigma^2[n]} P_{12}^-[n] \\ &\cdots\cdots\cdots\cdots \\ P_{1m}^+[n] &= P_{1m}^-[n] - \frac{P_{11}^-[n]}{P_{11}^-[n] + \sigma^2[n]} P_{1m}^-[n] \\ P_{22}^+[n] &= P_{22}^-[n] - \frac{P_{12}^-[n]}{P_{11}^-[n] + \sigma^2[n]} P_{22}^-[n] \\ &\cdots\cdots\cdots\cdots \\ P_{mm}^+[n] &= P_{mm}^-[n] - \frac{P_{1m}^-[n]}{P_{11}^-[n] + \sigma^2[n]} P_{mm}^-[n] \end{aligned}\right\} \tag{12}$$

Best Available Copy

Calculation of the elements of the matrix P^- is done according to extrapolation model of the signal:

$$\left.\begin{aligned} \hat{E}_1^-[n] &= a_{11}\hat{E}_1[n-1] + \ldots + a_{1m}\hat{E}_m[n-1] \\ \hat{E}_2^-[n] &= a_{21}\hat{E}_1[n-1] + \ldots + a_{2m}\hat{E}_m[n-1] \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ \hat{E}_m^-[n] &= a_{m1}\hat{E}_1[n-1] + \ldots + a_{mm}\hat{E}_m[n-1] \end{aligned}\right\} \quad (13)$$

on the basis of the following formulae of the probability theory:

$$P_{ij}^-[n] = M\{\hat{E}_i[n-1]\,\hat{E}_j[n-1]\} = f(P_{ij}^+[n-1]\,a_{ij}) \quad (14)$$

In each step of the procedure of the estimation algorithm, we do the following calculations: compute E[n] according to (13); form the measurement $Z_1[n]$; compute P_{ij}^- according to [14]; determine E[n] according to (11); and compute $P_{ij}^+[n]$ according to (12).

The procedure for defining the number of dynamic stereotypes was done on the basis of analysis of training test data of 36 trained operators. The OEF was computed for each test of each subject. Dynamic stereotypes were then selected on the basis of the OEF's. Dynamic stereotypes D_i were determined in a two-step procedure. In the first step the basic D_i were determined according to the following set of parameters: range of final OEF values and number of test values. In the second step, D_i were determined more precisely by a digital spectral analysis technique. First, normalized estimation of the integrated sample co-spectrum was determined, which measures the total in phase covariance between the two processes for all frequencies less than f_o:[12]

$$I_{12}(f_e) = \frac{2}{N S_1 S_2 \Delta t} \sum_{i=0}^{K} \bar{C}_{12}(f_i) \quad (15)$$

where

$$\bar{C}_{12}(f_i) = 2\{\hat{c}_{12}(0) + \sum_{K=1}^{\ell-1} \hat{c}_{12}(K)\, w(K) \cos(\pi i K/\bar{H})\}, \quad 0 \le i \le F \quad (16)$$

$$\hat{C}_{12}(K) = \frac{1}{2}\{C_{12}(K) + C_{12}(-K)\}, \quad 0 \le K \le L-1 \quad (17)$$

$$\hat{C}_{12}(K) = \frac{1}{N} \sum_{t=1}^{N-K} (E_{1t} - \bar{E}_1)(E_{2t+K} - \bar{E}_2), \quad 0 \le K \le L-1 \quad (18)$$

S_1 and S_2 are estimators of the standard deviations.

It is known that if the two processes are uncorrelated then $I_{12}(f_e)$ is exactly zero, but if the two processes are correlated $I_{12}(f_e)$ takes on non-zero values. Estimations of I_{mn} for a certain frequency $f_K = 1/2\Delta t$ were computed for all combinations from the set of functions $E_{op}(t) \subset E$. As a result, four basic dynamic stereotypes were determined for this task. For each D_i, the average OEF time history was simulated. On the basis of analysis, it was possible to estimate the influence of the numbers of competitors N and N_o on the relative decreasing time of stabilization of final OEF values. Figure 9a shows the result for eleven competitors and Fig. 9b for five. It is possible to see that decreasing the number N_o in the leader group (i.e., subject should initially be one of the top N competitors, $N_o = N$; then N-1, $N_o = N-1$; ... and finally he should finish first, $N_o = 1$), requires the subject to complete the task in steadily decreasing amounts of time. This involves increasing productivity on the part of the subject and simultaneously, the time required to stabilize the final OEF value decreases. Also the deviation of the values of $\Delta T/T$ is decreasing as the task becomes more complicated.

Training in an environment of competition was conducted with four simulated competitors. Display of competion indicated (Fig. 10) five bars, the heights of which proportionally indicated the current values of OEF of operator B and four simulated competitors (a); elapsed time of test (b); the percentage of task goal remaining (c); and number of the test run (d).

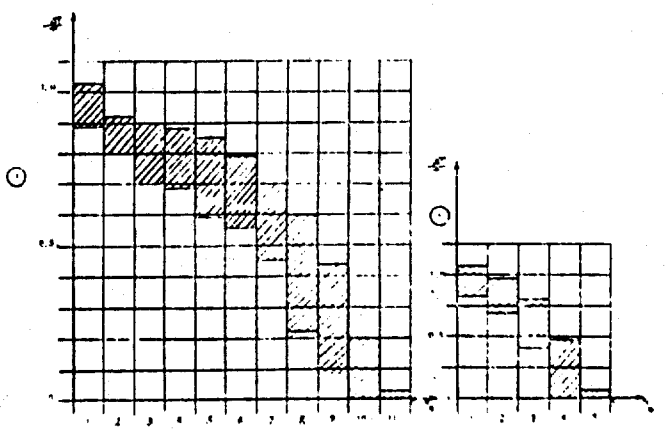


Fig. 9 Typical Relationship Between Productivity of Human Behavior and the Number of the Leader Group Computed for Eleven Competitors in 9(a) and for Five Competitors in 9(b)

Best Available Copy

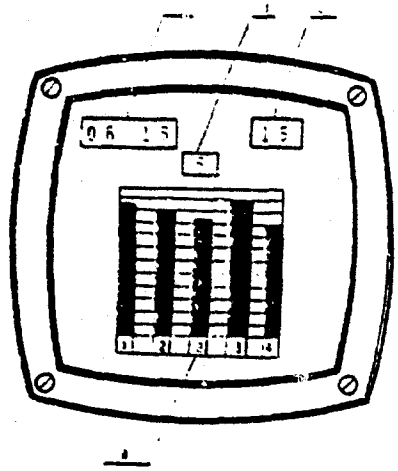


Fig. 10 Competition Display

The operator is matched to one of the four basic dynamic stereotypes on the basis of his final OEF values and the range of final OEF values of the basic dynamic dynamic stereotypes, as well as the parameter O_{E_i}:

$$O_{E_i} = \frac{\frac{1}{T}\sum_{j=1}^{N}[E_{op}(t_j) - \hat{E}_{D_i}(t_j)]\,\Delta t}{\sigma^2_{E_i}} \qquad (19)$$

where

$$\sigma^2_{E_i} = \frac{1}{T}\sum_{j=1}^{N}\{[E_{op}(t_j) - \hat{E}_{D_i}(t_j)] - \mu_{E_i}\}^2\,\Delta t \qquad (20)$$

and

$$\mu_{E_i} = \frac{1}{T}\sum_{j=1}^{N}\{E_{op}(t_j) - \hat{E}_{D_i}(t_j)\}^2\,\Delta t \qquad (21)$$

which shows the level of oscillations for $E_{op}(t)$ relative to $\hat{E}_{D_i}(t)$.

Figure 11 shows the typical time histories of the OEF for five channels of manual control: velocity of angular stability (ω_y, ω_z); fuel consumption rate (F_x, F_y); and relative velocity of approach $\dot{o}$. Column a contains test data without the environment of competition and data in column b was taken using the environment of competition. Both operators had taken 15 test runs. At the bottom of each column, values of OEF are shown integrated across all channel with respect to time.

In order to further determine the effectiveness of the training methodology, we compared the results obtained by training two populations of pilots whose previous professional competitiveness were considered very similar. One population A, was trained using the classic methodology; population B was trained by the environment of competition method. Figure 12 shows the average final OEF value for

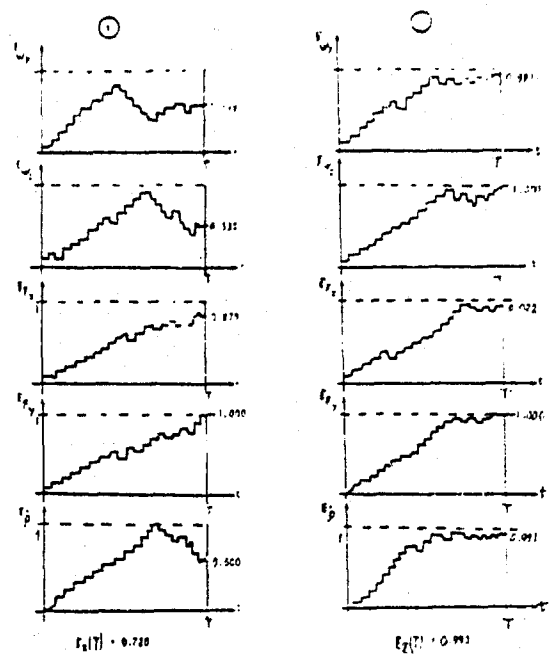


Fig. 11 Comparison of the Typical Histories of the OEF for Five Channels of Manual Control

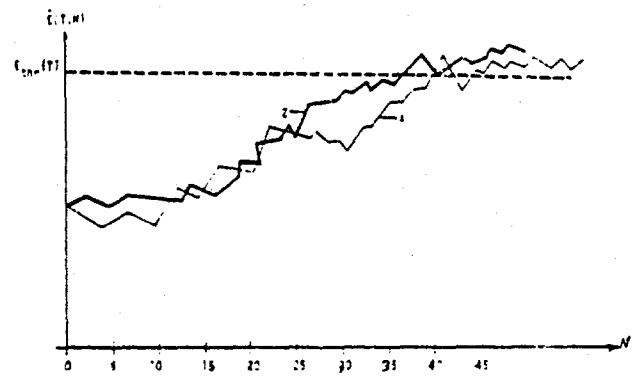


Fig. 12 Comparison of the Relationships Between the Average Final OEF Value and the Number of Test Runs

these subjects as a function of the number of test runs completed with competition methodology (2) and without competition (1).

Conclusion

In spite of some difference between subjects, both functions (Fig. 12) show us that in an environment of competition the threshold value $E_{thr}(\bar{T})$ for OEF can be reached with fewer test runs than by traditional methodology; and the stable value of the final OEF reached with the environment of competition was higher than by traditional methodology.

A methodology has been described, which when combined with manned-vehicle systems analysis can aid further investigations of human behavior in man-machine systems to increase their efficiency.

Best Available Copy

Further Work

In the future, we expect the following topics to be of interest: research in optimal training strategies based on the game theory; and reliability estimations of the results of pilot training.

Acknowledgements

The author wishes to express his sincere appreciation to Professor Thomas B. Sheridan (MIT), Mr. Michael Zeldin (Harvard University) for their helpful discussion and to Mr. Dana R. Yoerger (MIT) who helped in the transcription of this paper. Also the author's appreciation goes to Ms. Gwen Heistand and to Ms. Bonnie Beatty for their assistance in preparing this paper.

References

1. Sheridan, T.S., Ferrell, W.R. Man-Machine Systems. MIT Press, Cambridge, Massachusetts (1974).

2. Tarran, V.A., Lipchin, L., Orlov, M.G., ZlotniR, A. "A System for the Optimal Selection and Training of Crew Members." Author's license (certificate) no. 445055. Bulletin of Inventions and Discoveries. Moscow, USSR, No. 36, 1974. (Russian and English)

3. Tarran, V.A., Lipchin, L., Metkin, N.P., Korotkov, A.D. "Some Problems of Optimal Design in Simulation Technology." Proceedings All Union Conference On Information Systems Of Automatic Control And Operation. Moscow, USSR, 1972, pp. 105-106. (Russian)

4. Arhangelsky, B.P., Lipchin, L., Marchenko, V.V. "Some Problems In The Design of Autonomous Training Systems For Pilots." Proceedings All Union Conference On Information Systems Of Automatic Control And Operation. Moscow, USSR, 1972, pp. 106-107. (Russian)

5. Metkin, N.P., Mironov, V.M., Lipchin, L., Schustov, A.K. "Analysis of Accuracy of Formation of Some Statistics Characteristics Of The Simulator of Random Process." Proceedings VI Union Symposium On Methods Of Presentation And Analysis Of Some Random Processes And Fields. Erevan, USSR, 1973, Vol. 4, pp. 28-36. (Russian)

6. Scherbak, V.I., Lipchin, L., Hohlov, I.I. "Standardized Computer Procedures of Man-Machine Engineering Design." Express Standard (Quality, Standardization, Methodology). Moscow, USSR, No. 53 (539), 1975, pp. 10-13. (Russian and English)

7. Orlov, M.G., Tarran, V.A., Lipchin, L. "A Methodology For Selection Training And Optimal Utilization Of Spacecraft Crew In Atmosphere Of Competition." Proceedings All Union Conference On Computer Science In Psychology. Gorky, USSR, 1977, pp. 83-87. (Russian)

8. Lipchin, L. Design And Manufacturing Of Navigation Computing Systems. Mashinostroenie, Moscow, USSR, 1976. (Russian)

9. Tarran, V.A. Ergonomic Control Systems. Mashinostroenie, Moscow, USSR, 1976. (Russian)

10. Artukhovsky, O.A., Lipchin, L., Orlov, M.G. "Kalman's Filter In Electrocardiographic Investigations." Mathematical Medico-Biological Processing. Nauka, Moscow, USSR, 1976, pp. 222-226. (Russian)

11. Bush, R.R., Mosteller, F. Stochastic Models For Learning. Wiley, New York, 1955.

12. Jenkins, G.M., Watts, D.G. Spectral Analysis And Its Applications. Holden-Day, 1969.

13. Lerner, M., Lipchin, L. Reliability Prognosis Of Electrical Capacitors Subjected To A Non-Sinusoidal Voltage. In Proceedings 17th Annual Spring Reliability Seminar, Mass., April, 1979.

14. Harper, R.P., Cooper, G.E. A Revised Pilot Rating Scale For The Evaluation of Handling Qualities. Cornell Aeronautical Laboratory Report No. 153, September 1966.

15. McDonnell, J.D. "An Application Of Measurement Methods To Improve The Quantitative Nature Of Pilot Rating Scales." IEEE Trans. Man-Machine Systems, Vol. MMS-10, No.3, pp. 81-92, September 1969.

16. Moorhouse, D.J. "The History And Future Of The U.S. Military Flying Qualities Specifications." American Institute of Aeronautics And Astronautics Paper 70-0402, Seventeenth Aerospace Sciences Meeting, New Orleans, Louisiana, January 1979.

17. Onstott, E.D., Faulkner, W.H. "Definition And Interpretation Of Flying Qualities." NOR 79-22, Northrop Corporation, Hawthorne, California, March 1979.

Best Available Copy

SESSION 3: HUMAN DYNAMICS MODELING

Moderator: Dr. Carroll Day, 6570 Aerospace Medical Research Lab

EFFECTS OF MUSCLE VIBRATION AND JOINT OSCILLATION ON HUMAN MOTOR MECHANISMS

Gyan C. Agarwal
Gerald L. Gottlieb

College of Engineering
University of Illinois at Chicago Circle
Chicago, Illinois 60680

and

Department of Physiology
Rush Presbyterian-St. Luke's Medical Center
Chicago, Illinois 60612

ABSTRACT

Vibration at frequencies above 50 Hz applied to the tendon of the extensor muscles of the ankle joint produce the tonic vibration reflex (TVR) which increases when the vibration frequency is increased. The TVR affects a joint's mechanical response to both sinusoidal and random oscillations in a manner similar to that seen with tonic voluntary contraction. Vibration also inhibits the myotatic component of the stretch reflex measured in the EMG. The degree of inhibition is proportional to the vibration frequency. This is in contrast to the facilitation of the myotatic reflex produced by tonic voluntary contraction. Vibration does not influence the post-myotatic component of the stretch reflex. The selective influence of vibration on myotatic and post-myotatic responses suggests different anatomical and functional organization underlying these responses.

INTRODUCTION

The suppression of the patella reflex in human subjects seated on a vibrating platform was observed some forty years back by Coermann (1938, cited by Goldman, 1948). Interest in this effect of vibration was renewed by the studies of Hagbarth and Eklund (1966) and DeGail, Lance, and Neilson (1966). The tonic vibration reflex (TVR) and the suppression of both tendon-jerk and Hoffmann reflexes were observed in each of these studies. Numerous studies

since have shown that vibration of the limb in animals as well as humans activates both monosynaptic and polysynaptic spinal loops (Hagbarth 1973; Lance, Burke & Andrews, 1973; Matthews, 1972).

The exact mechanisms by which vibration suppresses monosynaptic reflexes are not fully resolved. Various mechanisms suggested include peripheral occlusion of Ia afferents, presynaptic inhibition of Ia inputs at the motoneuron level, fusimotor driven changes in spindle sensitivity, and reciprocal inhibition of motoneurons due to activation of primary endings of antagonist muscle groups (Barnes & Pompeiano, 1970; Fromm & Noth, 1976; Gillies et. al., 1969; Homma, et. al., 1975; Dindar& Verrier, 1975). Although most studies have shown inhibition of the monosynaptic reflexes, exceptions have been observed. Godaux & Desmedt (1975) have shown that 100 Hz vibration of the human masseter muscle potentiates both the masseter tendon-jerk and H-reflex. Agarwal & Gottlieb (1976) found the effect of vibration on the achilles tendon jerk to be frequency dependent. The tendon jerk was inhibited at 60 and 100 Hz, but was facilitated at 160 and 200 Hz vibration. The H-reflex was also inhibited at 60 and 100 Hz, slightly inhibited at 160 Hz and unchanged at 200 Hz. Desmedt & Godaux (1978) have shown that the inhibitory effects of vibration increase with amplitude, but decrease when the vibration frequency is increased. The paradox that on one hand maintained vibration activates a tonic contraction (the TVR) while on the other hand simultaneously inhibits the phasic reflexes is yet to be adequately explained.

The TVR produces involuntary movements when the joint is free to move and in an isometric situation it produces illusions of movements (Goodwin, McCloskey & Matthews, 1972; McCloskey, 1973; Craske, 1977). Jaeger, Agarwal & Gottlieb (1979) have shown that the vibration at 100 Hz applied to the muscle tendons of soleus and anterior tibial muscles,increases both the simple and choice reaction times in a discrete tracking task. This effect is presumably due to altered peripheral feedback which influences motor command decision making.

Hendrie and Lee (1978) have shown selective effects of vibration at 120 Hz applied to the human forearm on the M_1 and M_2 - M_3 responses recorded from the wrist flexor muscles. A marked suppression of M_1 was observed, but vibration did not significantly change the M_2 - M_3 component of the response.

In this paper we will present the results of experiments with vibration applied to the tendon of the extensor muscles of the ankle joint in normal human subjects, on the mechanical parameters of the joint as well as its selective effect on spinal mechanisms.

METHODS

Experiments were done on six normal, adult, human subjects. A subject sat in a chair with the right foot strapped to a footplate which could rotate about a horizontal, dorsal-plantar axis through the medial malleolus. A digital computer (General Automation SPC-16/65) recorded foot angle (θ) and torque (T) (sampling rate 250/s) and rectified, filtered EMG (sampling rate 500/s) from the soleus muscle (SM) and anterior tibial muscle (AT) using disk surface electrodes. The layout of the apparatus is shown in Figure 1 and a more complete description is available in Agarwal & Gottlieb (1977a). Vibration was applied to the Achilles tendon using a Hagbarth-type vibrator (TRV-Vibrator, Model #TMT-18, Heiwa Electronic Industrial Corp., Japan) attached with surgical tape. Vibrator frequencies were selected at 50, 75, 100, 125, or 150 Hz.

The plate could be rotated by a DC torque motor (Inertial Motors Corp. #06-24) via a gearbelt and pulley system for torque amplification. Constant tension springs were also used to counterbalance gravitational torque. With the subject completely relaxed, the resulting joint position (approximately 90 degrees between the foot and the tibia) was defined as the reference position of the ankle joint. A dual beam oscilloscope provided the subject with visual feedback of his foot angle on one channel and the reference position on the other.

Experiment 1: High frequency vibration (50-150 Hz) produces a tonic vibration reflex (TVR) contracting the gastrocnemius-soleus muscle. The resulting torque due to the TVR was counterbalanced by applying an appropriate level of bias torque with the motor so that the subject would maintain the reference position without voluntary effort. Sinusoidal torques in the frequency range of 3 to 12 Hz were then superimposed on the bias torque level. The experiment was repeated without vibration but with the same level of bias torque which required tonic voluntary contraction of the gastrocnemius-soleus muscles.

Fourier series coefficients were computed from the torque and angular rotation data to calculate the effective compliance of the ankle joint. (For details see Agarwal & Gottlieb, 1977a; Gottlieb, Agarwal, & Penn, 1978).

The compliance data, presented in Fig. 4 and 6, can be approximated by a linear mechanical system consisting of a moment of inertia (J), a viscous element (B), and an elastic element (K). Such a system is described by a linear differential equation:

$$T = J \frac{d\Theta}{dt^2} + B \frac{d\Theta}{dt} + K\Theta \qquad (1)$$

where T and Θ are the torque and angular rotation. In Laplace transform notation the complex compliance is given by:

$$C(S) = \frac{\Theta}{T} = \frac{1}{JS^2 + BS + K} \qquad (2)$$

where S is the complex frequency. The solid lines in Fig. 4 and 6 represent the 'best' fit of equation (2) to the data.

Experiment 2: A band-limited gaussian (0-15 Hz) signal was pre-recorded from a noise generator. These time-varying signals were superimposed on a bias torque level as in Experiment 1. The input was applied for 30 sec or more and the data continuously recorded on a digital tape.

The data was analyzed by computing the auto- and cross-power spectra of the angle and torque records using 4.096 sec data records (1024 points) and cosine taper with 2.048 sec of overlap between successive records. Transfer functions and coherence were computed (Agarwal & Gottlieb, 1977b; Gottlieb & Agarwal, 1978). Figure 6 shows the compliance magnitude and phase as calculated from the complex transfer function. The coherence function was calculated to check the linearity assumption. Average coherence values are given in Table 1.

Experiment 3: In the third experiment, pulses of torque lasting one second were applied to dorsiflex the foot. The instructions to the subject were to react as quickly as possible in opposition to the motor pulse and restore the foot to its original position. For details of the methods, see Gottlieb and Agarwal (1979).

Three to nine pulse amplitudes were used in most experiments. Pulse amplitudes and interstimulus intervals were in pseudorandom sequence. Ten responses at each amplitude level were averaged and graphed as shown in Figures 7 and 8.

Visual inspection of the EMG records determined the time interval over which a particular EMG response occurred and the area under each record for that interval was computed.

To compensate for shifts in the base line produced by tonic voluntary contraction (or due to TVR), we computed the average value of the EMG over the 50 ms interval prior to the torque stimulus and subtracted this from the EMG in the integration interval.

The rate of joint rotation and hence muscle stretch changes

rapidly upon application of a torque pulse. The average velocity in 16 to 24 ms interval after torque onset was used as the independent variable in graphing integrated EMG response versus stimulus intensity. The experiment was done without vibration and at three vibration frequencies (50, 100, and 150 Hz). When vibration was applied, the resultant TVR torque was counter-balanced by the motor bias torque to maintain the reference position. For comparison, experiments were also done without vibration with bias torques requiring comparable tonic voluntary contraction of gastrocnemius-soleus muscle.

RESULTS

Sinusoidal Oscillation

Figure 2 shows the first 2 s of oscillation due to a 5 Hz sinusoidal torque input without vibration and with vibration at 50, 100, and 150 Hz. The input torque amplitude is the same in all four cases. The amplitude of the angular rotation decreases with increasing vibration frequency indicating greater stiffness of the joint. There is very little EMG activity in the AT. For the SM, note the different EMG scale for no vibration case. There is a reduction of about 20% in the peak to peak amplitude of rotation with 50 Hz vibration, while the soleus EMG response is reduced by a factor of 2.

The width of the SM EMG burst is increased with vibration as shown in the two-cycle averages in Figure 3. In these averages, each horizontal trace represents the average response at a single frequency with 3 Hz in the back-ground and 12 Hz in the foreground. They show a clear reduction in the amplitude of angular rotation with increasing vibration frequency. In the SM, there is a significant decrease in the starting time of the EMG pulse with respect to the most plantar point of rotation as the vibration frequency is increased.

Figure 4 shows the compliance of the ankle joint and its phase angle as a function of the input frequency. The solid lines are the best linear second-order model fit. The mechanical parameters of the model are shown in Table 1. As the vibration frequency is increased, the bias torque needed to counterbalance the TVR also increases proportionately. The moment of inertia and the viscosity remain relatively constant but the joint stiffness increases by nearly a factor of two from no vibration to 150 Hz vibration. The resonant frequency increases from 5.9 to 7.3 Hz.

Random Torque Input

Figure 5 shows a response to a bandlimited (0-15 Hz) Gaussian torque input with no vibration, 150 Hz vibration, and with no vibration and a voluntary contraction equivalent to the previous TVR.

Figure 6 shows the compliance and phase angle. The parameters of the second order model are shown in Table 1. Again the moment of inertia and viscosity are nearly constant while the stiffness increases with vibration frequency. The effect of vibration on stiffness is exactly the same as with voluntary contraction against a bias torque without vibration. The last column of this table shows the average coherence value over the frequency range of 2 to 15 Hz used for curve fitting.

Stretch Response

Figures 7 and 8 show the results of the third experiment for two different subjects. Part (a) shows averaged responses of angular rotation and the SM EMG. The step torque is applied 50 ms after the start of the record. In Figure 7a the two cases are without vibration (top) and with 50 Hz vibration (bottom). In Figure 8a (bottom) vibration is at 100 Hz. In both subjects, the first component of the EMG response between 30 and 100 ms after the torque application (this corresponds to 80 and 150 ms in these plots) is significantly reduced with vibration but the second component is nearly unaltered. These responses are quantitatively compared by plotting the integrated EMG in appropriate intervals versus the stretch velocity as shown in parts (b) and (c). The straight lines are first order regression curves. The slopes of these lines are defined as the response gain.

Table 2 shows the gain, the correlation coefficient and the velocity threshold (intercept of the regression line on the abscissa) for the myotatic component (before 100ms) and the post-myotatic component (after 100 ms). Responses without vibration were done at the start of the experiment (x-points) and repeated after the vibration inputs (v-points). For subject GCA, the myotatic reflex gain is reduced by a factor of 10 between the no-vibration and 150 Hz vibration. For subject BXK, the gain is reduced by a factor of 3 with 100 Hz vibration. It should be emphasized that the tonic voluntary contraction of the soleus muscle increases the myotatic reflex gain (Gottlieb and Agarwal, 1979). The last line in the table for subject BXK represents the condition of tonic voluntary contraction without vibration which demonstrates this increase.

The gain of the post-myotatic responses, 120 to 250 ms in Fig. 7(c) and 130 to 250 ms in Fig. 8(c), is not significantly influenced by vibration. Just as with the earlier myotatic reflex, this response is highly and linearly correlated with the rate of ankle rotation.

Sinusoidal & Random Torque Inputs

The visco-elastic parameters of the ankle joint have been shown to be linear functions of the level of muscle contraction. (Agarwal & Gottlieb, 1977b; Gottlieb & Agarwal, 1978). The vibration produced TVR contraction of the muscle changes its mechanical properties in a manner not distinguishable from voluntary contraction. The co-variation of ankle joint stiffness with joint torque is the same whether that joint torque is produced by voluntary contraction or the TVR (table 1).

In an earlier study (Agarwal and Gottlieb, (1977a) we have shown that the resonance in compliance near 6 Hz appears to be in part a manifestation of the stretch reflex arc (see also Joyce et. al., 1974; Goodwin et. al., 1978). Vibration reduces the tendon-jerk response, the Hoffmann reflex and the stretch reflex, implying that the spindle afferent input (Ia) is not as effective on the alpha motoneuron. These reductions have been attributed to presynaptic inhibition of the primary afferent-alpha motoneuron arc (Gillies et. al., 1969, Barnes and Pompeiano, 1970, Delwaide, 1973, Dindar & Verrier, 1975).

In the present experiments it is clear that the repeated stretch of the soleus by sinusoidal torques overcomes the vibration induced inhibition of the Ia-alpha MN arc. In Fig. 2, the SM EMG builds up in amplitude after each cycle and in the averaged response in Fig. 3, the SM EMG amplitudes are of the same order at all vibration frequencies. With random torque inputs, (Fig. 5) vibration at 150 Hz produces a strong TVR response over which the stretch induced EMGs are superimposed with apparently reduced magnitudes.

Step Inputs

The myotatic component of the stretch reflex in the human soleus muscle is highly and linearly correlated with the rate of muscle stretch (Gottlieb & Agarwal, 1979). The slope of this response curve is proportional to the level of tonic voluntary activation. The response curves in figures 7 and 8 show a linear relationship to the applied stretch. However, the tonic muscle activity of the TVR inhibits the myotatic response. This indicates that the motoneuron pool is activated in a different manner by the TVR and by voluntary tonic contraction.

The post-myotatic response is not influenced by vibration. We have also found that the tonic contraction prior to stretch input

has only a slight effect on its magnitude (Gottlieb & Agarwal, manuscript in preparation). Regardless of the mechanisms involved in the suppression of the myotatic component, the invariance of the post-myotatic response suggests that this response is mediated by peripheral afferent pathways in addition to the Ia afferents which are responsible for most or all of the myotatic component of the response. The step torque inputs undoubtedly excite group Ib and secondary spindle afferents (II) as well as joint and cutaneous receptors. Burke et. al. (1976) have shown in man that vibration is not as selective a stimulus for spindle primaries as had been believed from animal experiments.

Even if presynaptic inhibition at Ia terminals is the predominant mechanism for myotatic reflex suppression, it may not be operative on the central projections of muscle afferents to supraspinal centers (Hendrie & Lee, 1978). It has also been suggested that separate subpopulations of the spinal motoneuron pool are responsible for the M_1, M_2 and M_3 peaks of the EMG response (Tatton, et. al., 1978). These may be differentially susceptible to vibration as well.

CONCLUSIONS

The tonic vibration reflex affects a joint's response to sinusoidal oscillation in the same facilitatory manner that is seen with tonic voluntary contractions. Although the myotatic reflex is suppressed by vibration, repeated stretches of sinusoidal oscillation produce an average EMG response which is not different in magnitude from the no vibration case. Either polysynaptic mechanisms at the spinal cord level or mechanisms involving higher centers (and possibly both) are able to overcome the inhibitory mechanisms at Ia-alpha motoneuron level in producing stretch evoked resonance near 6 Hz.

These results also indicate that the post-myotatic EMG responses to limb perturbation are not only different in their latency but also in their functional dependence upon peripheral influences.

ACKNOWLEDGEMENT

This work was supported by NSF grant ENG - 7608754 and NIH grants NS - 00196 and NS - 12877.

REFERENCES

1. Agarwal, G.C. & Gottlieb, G.L. (1976). Effects of vibration on human spinal reflexes. In The Motor System: Neurophysiology and Muscle Mechanisms, M. Shahani (Editor), 181-188, Elsevier.

2. Agarwal, G.C. & Gottlieb, G.L. (1977a). Oscillation of the human ankle joint in response to applied sinusoidal torque on the foot. J. Physiol (London), 268, 151-176.

3. Agarwal, G.C. & Gottlieb, G.L. (1977b). Compliance of the human ankle joint. ASME J. of Biomechanical Engineering, 99, 166-170.

4. Barnes, C.D. & Pompeiano, O. (1970). Presynaptic inhibition of extensor monosynaptic reflex by Ia afferents from flexors. Brain Res. 18, 380-383.

5. Burke, D., Hagbarth, K.E., Lofstedt, L., & Wallin, B.G. (1976). The responses of human muscle spindle endings to vibration of non-contracting muscles. J. Physiol., 261, 673-693.

6. Burke, D. & Schiller, H.H. (1976). Discharge pattern of single motor units in the tonic vibration reflex of human triceps surae. J. Neurol. Neurosurg. Phychiat., 39, 729-741.

7. Craske, B. (1977). Perception of impossible limb positions induced by tendon vibration. Science, 196, 71-73.

8. De Gail, P., Lance, J.W. & Neilson, P.D. (1966). Differential effects on tonic and phasic reflex mechanisms produced by vibration of muscles in man. J. Neurol. Neurosurg. Psychiat., 29, 1-11.

9. Delwaide, P.J. (1973). Human monosynaptic reflexes and presynaptic inhibition. In New Developments in Electromyography and Clinical Neurophysiology, Vol. 3, J.E. Desmedt (Editor), pp. 508-522, Basel: Karger.

10. Desmedt, J.E. & Godaux, E. (1978). Mechanism of the vibration paradox: Excitatory and inhibitory effects of tendon vibration on single soleus muscle motor units in man. J. Physiology, 285, 197-207.

11. Dindar, F. & Verrier, M. (1975). Studies on the receptor responsible for vibration induced inhibition of monosynaptic reflexes in man. J. Neurol. Neurosurg. Psychiat. 38, 155-160.

12. Fromm, C. & Noth, J. (1976). Reflex responses of gamma motoneurones to vibration of the muscle they innervate. J. Physiol., 256, 117-136.

13. Gillies, J.D., Lance, J.W., Neilson, P.D., & Tassinari, C.A. (1969). Presynaptic inhibition of the monosynaptic reflex by vibration. J. Physiol. 205, 320-339.

14. Godaux, E. & Desmedt, J.E. (1975). Human masseter muscle: H - and tendon reflexes. Archives Neurol. 32, 229-234.

15. Goldman, D.E. (1948). Effect of mechanical vibration on the patella reflex in the cat. Amer. J. Physiology, 155, 78-81.

16. Goodwin, G.M., Hoffman, D. & Luschei, E.S. (1978). The strength of the reflex response to sinusoidal stretch of money jaw closing muscles during voluntary contraction. J. Physiol. (London), 279. 81-111.

17. Goodwin, G.M., McCloskey, D.I. & Matthews, P.B.C. (1972). The contribution of muscle afferents to kinaesthesia shown by vibration induced illusions of movement and by the effects of paralysing joint afferents. Brain, 95, 705-748.

18. Gottlieb, G.L. & Agarwal, G.C. (1978). Dependence of human ankle compliance on joint angle. J. Biomech., 11, 177-181.

19. Gottlieb, G.L. & Agarwal, G.C. (1979). The response to sudden torques about the ankle in man: The myotatic reflex. J. Neurophysiol., 42: 91-106.

20. Gottlieb, G.L., Agarwal, G.C., & Penn, R. (1978). Sinusoidal oscillation of the ankle as a means of evaluating the spastic patient. J. Neurol. Neurosurg. Psychiat., 41, 32-39.

21. Hagbarth, K.E. (1973). The effect of muscle vibration in normal man and in patients with motor disorders. In New Developments in Electromyography and Clinical Neurophysiology, J.E. Desmedt (Editor), Vol 3, Karger, Basel, 428-443.

22. Hagbarth, K.E. & Eklund, G. (1966). Motor effects of vibratory muscle stimuli in man. In Nobel Symposium I - Muscular Afferents and Motor Control, R. Granit (Editor), Almquist & Wiksell, Stockholm, 177-186.

23. Hagbarth, K.E., Hellsing, G. & Lofstedt, L. (1976). TVR and vibration-induced timing of motor impulses in the human jaw elevator muscles. J. Neurol. Neurosurg. Psychiat., 39, 719-728.

24. Hendrie, A. & Lee, R.G. (1978). Selective effects of vibration on human spinal and long-loop reflexes. Brain Res., 157, 369-375.

25. Homma, S., Mizote, M. & Watanabe, S. (1975). Participation of mono - and polysynaptic transmission during tonic activation of the stretch reflex arcs. Jap. J. Physiol., 25, 135-146.

26. Jaeger, R.J., Agarwal, G.C., & Gottlieb, A.L. (1979). Directional errors of movement and their correction in a discreate tracking task. J. Motor Behavior. (In press). See also Proceedings of the Fourteenth Annual Conference on Manual Control, pp.3-18, 1978.

27. Joyce, G.C., Rack, P.M.H. & Ross, H.F. (1973). The forces generated at the human elbow joint in response to imposed sinusoidal movements of the forearm. J. Physiol. (London), 240, 351-374.

28. Lance, J.W., Burke, D. & Andrews, C.J. (1973). The reflex effects of muscle vibration. In New Developments in Electromyography and Clinical Neurophysiology, J.E. Desmedt (Editor), vol. 3, Krager, Basel, 444-462.

29. Matthews, P.B.C. (1972). Mammalian Muscle Receptors and Their Control Actions. E. Arnold Publishers, New York.

30. McCloskey, D.I. (1973). Differences between the senses of movement and position shown by the effects of loading and vibration in man. Brain Research, 63, 119-131.

31. Tatton, W.G., Bawa, P., Bruce, I.C., & Lee, R.G. (1978). Long loop reflexes in monkeys: An interpretative base for human reflexes. In Cerebral Motor Control in Man: Long Loop Mechanisms, Prog. Clin. Neurophusiology, Vol. 4, J.E. Desmedt (Editor), pp. 229-245, Basel: Krager.

TABLE 1: Mechanical Parameters of the Ankle Joint
(Subject GCA)

Input Type (Date)	Vibration Frequency (Hz)	Bias Torque Kg.m.	J_2 Nms^2/rad	B Nms/rad	K Nm/rad	ζ	f_n	Average Coherence
Sinusoidal	-	0	0.024	0.201	32.3	0.115	5.9	-
(6/22/78)	50	0.35	0.027	0.201	41.2	0.095	6.2	-
	100	0.65	0.028	0.249	53.9	0.100	6.9	-
	150	1.00	0.030	0.306	62.6	0.112	7.3	-
Gaussian Torque 0-15 Hz Band-limited (2/2/79)	-	0	0.027	0.56	43.2	0.25	6.3	0.84
	50	0.68	0.028	0.42	45.2	0.19	6.4	0.85
	75	0.79	0.029	0.46	46.4	0.20	6.4	0.86
	100	0.88	0.028	0.46	46.4	0.20	6.5	0.89
	125	1.29	0.029	0.44	51.4	0.18	6.7	0.90
	150	1.46	0.029	0.48	54.0	0.19	6.9	0.90
	-	0	0.027	0.58	36.0	0.29	5.8	0.82
	-	0.68	0.028	0.44	46.0	0.19	6.4	0.85
	-	0.73	0.027	0.46	49.4	0.20	6.8	0.83
	-	0.88	0.028	0.44	47.0	0.20	6.5	0.85
	-	1.26	0.027	0.40	50.0	0.17	6.8	0.81
	-	1.46	0.028	0.46	53.6	0.17	7.0	0.84

TABLE 2: Reflex Regression Analysis

Subject	Vibration Frequency	Bias Kg.M.	Plot Symbol	Myotatic			Post-Myotatic		
				Gain	r	V_T	Gain	r	V_T
GCA	0	0.14	X	0.267	0.966	16	0.243	0.941	-8
	0	0.05	V	0.237	0.941	3	0.205	0.981	14
	50	0.06	*	0.145	0.935	41	0.149	0.924	-78
	100	0.70	□	0.070	0.999	62	0.190	0.853	-42
	150	0.64	△	0.019	0.370	-115	0.264	0.942	-15
BXK	0	0.04	X	0.095	0.998	20	0.111	0.930	-27
	0	-0.12	V	0.085	0.986	15	0.105	0.996	-49
	50	0.33	*	0.063	0.974	19	0.129	0.989	-10
	100	0.15	□	0.029	0.920	13	0.121	0.903	-38
	0	0.30	△	0.115	0.984	15	0.126	0.955	-10

Gain (μ.V.S./deg/s), velocity threshold V_T (deg/s), and Correlation Coefficient (r) between stretch velocity and integrated SM EMG. The subjects were instructed to resist to go back to the base line as quickly as possible. These data correspond to Figures 7 and 8. The bias is the average value of the torque measured just prior to the step torque input. When vibration input was applied, the TVR was counterbalanced by applying appropriate bias torque with the motor to maintain reference position. The gain corresponds to the slope and the threshold to the abscissa intercept of the linear regression line.

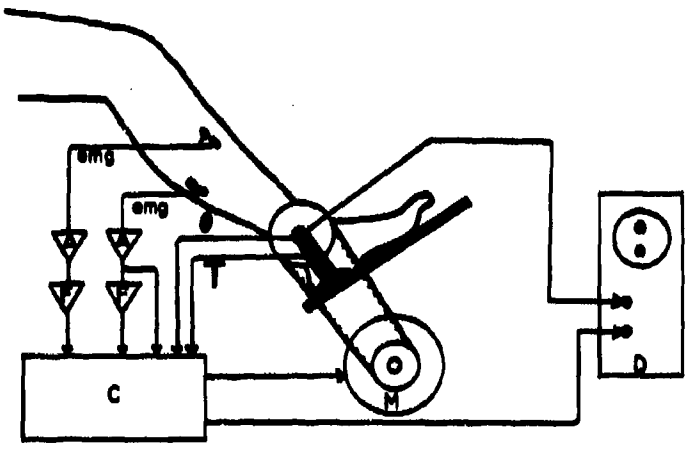


Fig. 1: Experimental motor driven footplate. The computer (C) records the amplified (A), filtered (F) EMGs as well as the ankle joint angle (θ) and applied torque (T). It drives the torque motor (M). Display dual beam oscilloscope (D) shows the subject a signal indicative of foot angle and the reference position.

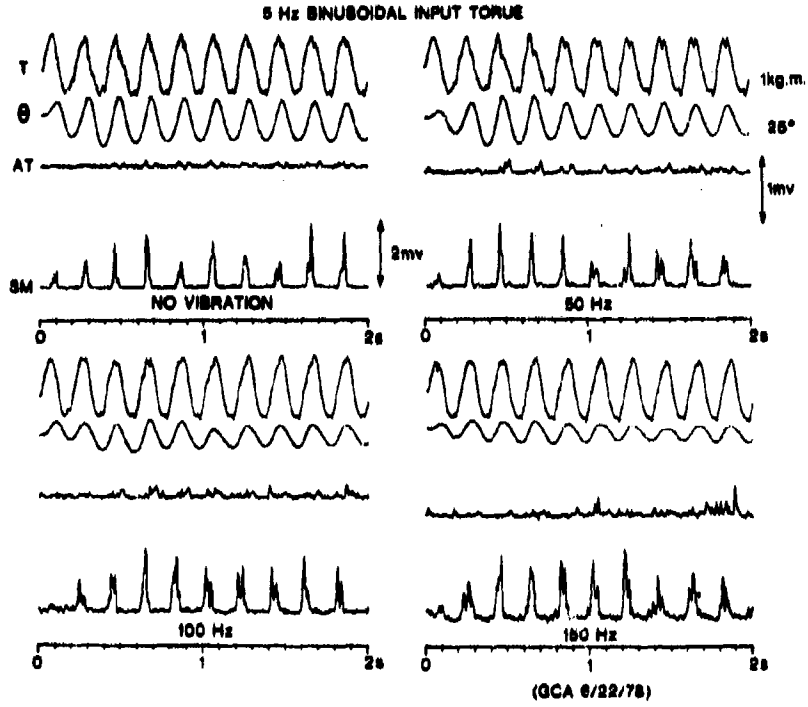


Fig. 2: Oscillation of the foot to an applied 5 Hz sinusoidal torque input under conditions of no vibration and achilles tendon vibrations at 50, 100, and 150 Hz. The TVR was counterbalanced by the torque motor. The four traces in each case are torque (T) in Kg. m., angular rotation (θ) in degrees, EMGs of anterior tibial (AT) and soleus muscles (SM). All EMG scales are 1 ms except SM EMG under no vibration condition.

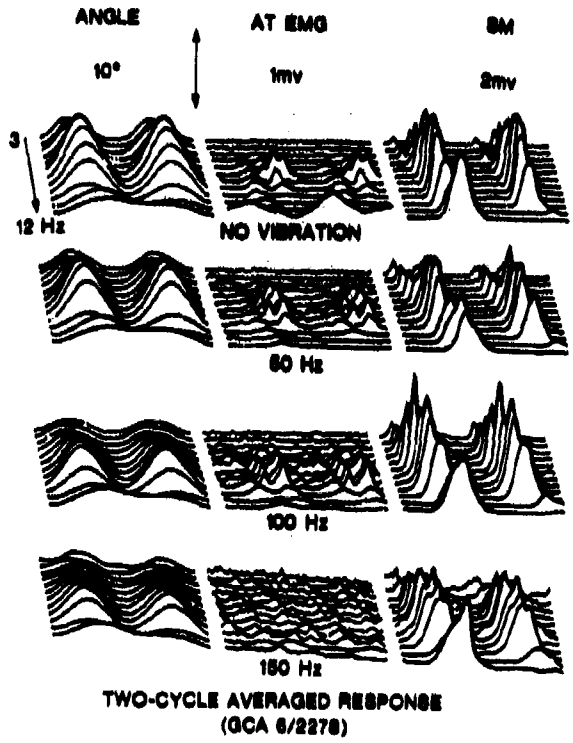


Fig. 3: Two-cycle averaged foot angle and EMGs from the anterior tibial (AT) and soleus muscles (SM) at different drive frequencies and under conditions of no vibration and at 50, 100, and 150 Hz soleus tendon vibration. The TVR was counterbalanced. The frequencies from background to foreground are 3, 3.5, 4, 4.5, 5, 5.5, 6, 6.5, 7, 8, 10, and 12 Hz.

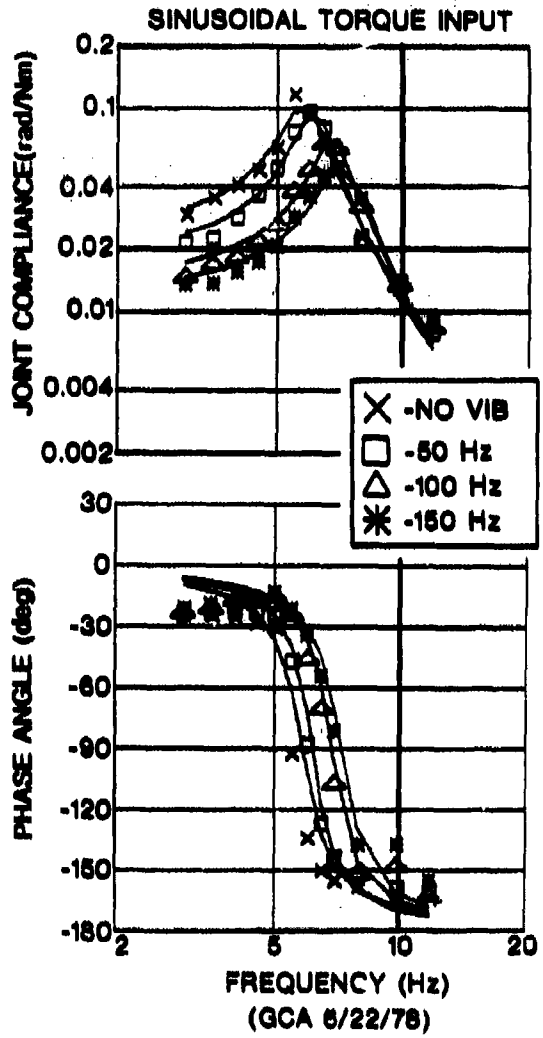


Fig. 4: Compliance of the ankle joint for the data shown in Fig.2. The solid lines in the compliance and phase data represent the best second order model fit. The mechanical parameter values are given in Table 1.

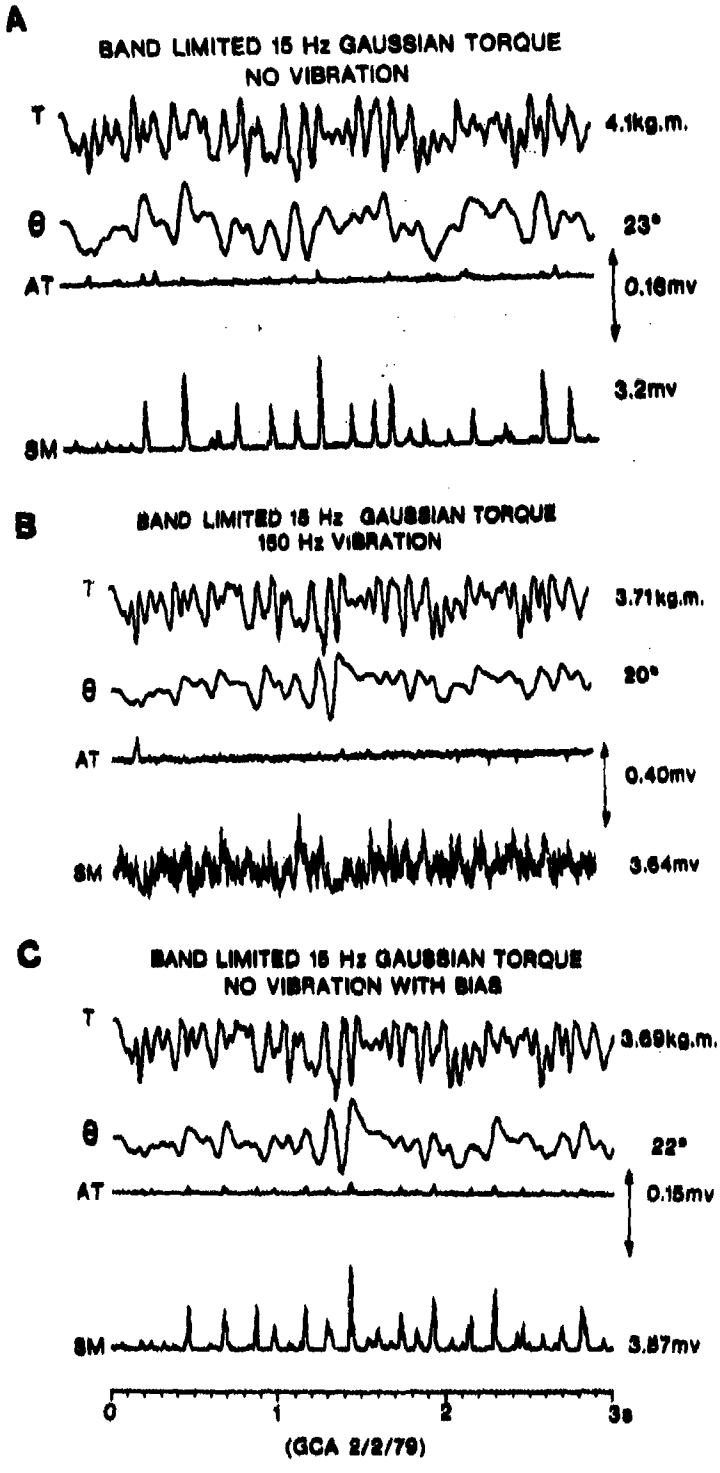


Fig. 5: Oscillation of the foot to an applied band limited (0-15 Hz) gaussian torque under conditions of no vibration (relaxed), 150 Hz vibration with TVR counterbalanced with motor bias, and no vibration with motor bias equal to the TVR at 150 Hz vibration. First three seconds of data are shown.

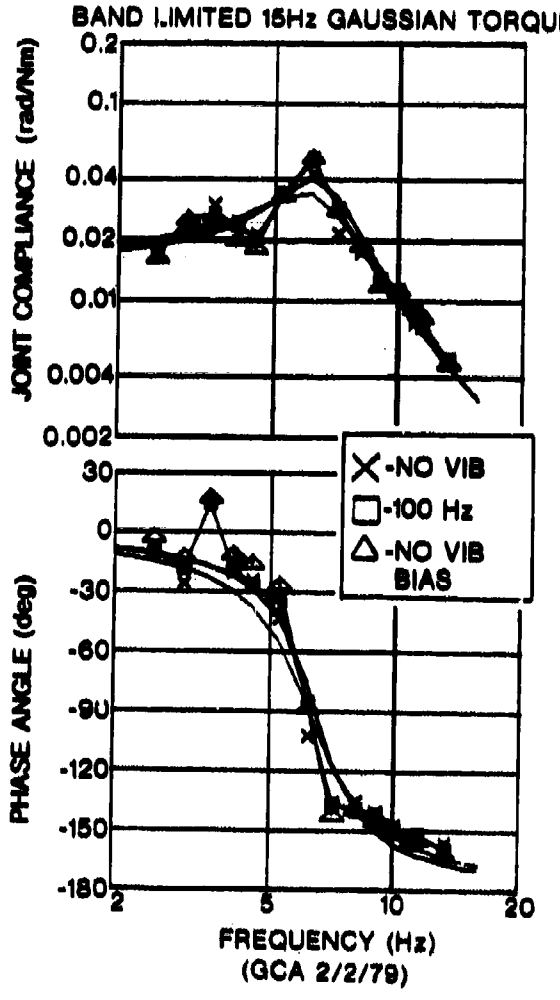


Fig. 6: Compliance of the ankle joint measured using random input torque signal. The model parameters are given in Table 1.

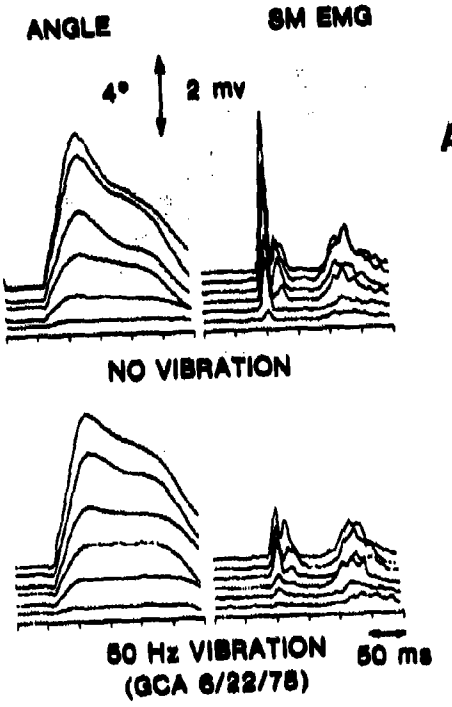


Fig. 7: a) Angular displacement and rectified, filtered SM EMG following sudden forced dorsiflexion of the foot. Torque is applied 50 ms after the start of each record. Each trace is the average of 10 responses, and responses to progressively stronger torque pulses are plotted from foreground to background. The upper group is under no vibration condition and the lower group under 50 Hz soleus tendon vibration condition. Dorsiflexion is plotted upwards.

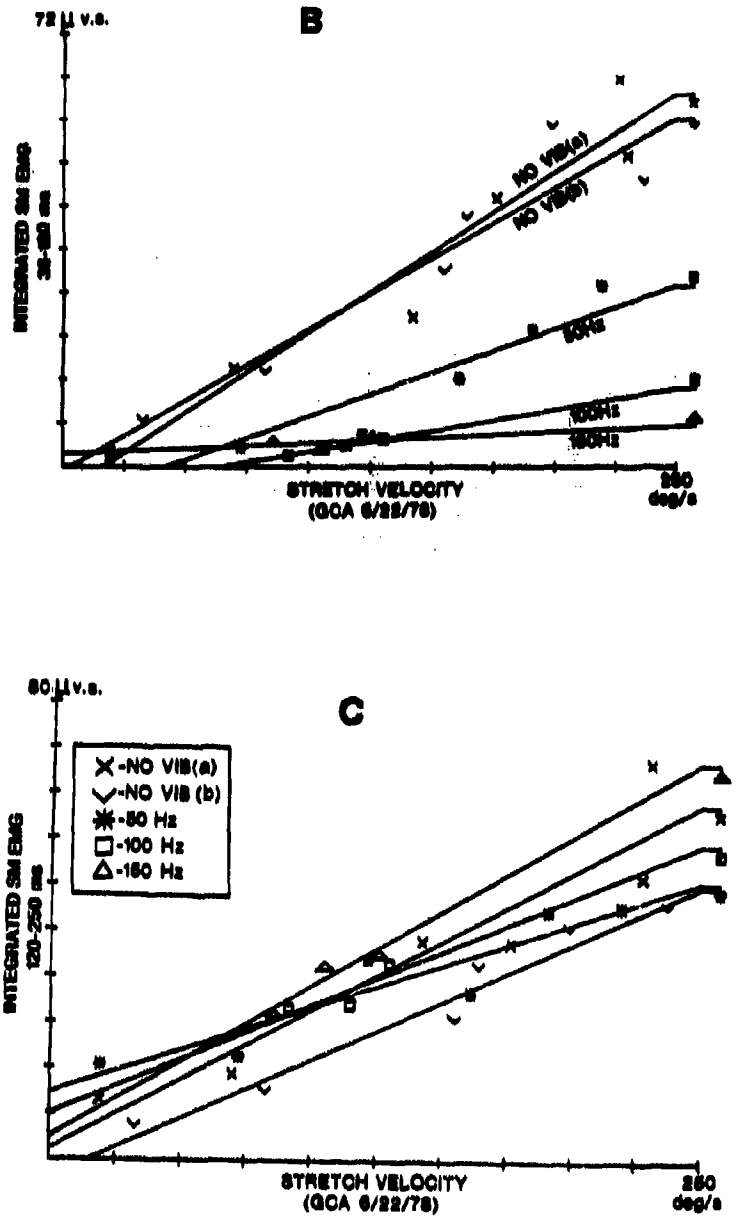


Fig. 7: b) Integrated SM EMG plotted against the angular velocity of forced dorsiflexion. The EMG was integrated over the interval 36-100 ms. The solid lines are linear regression curves and the legend next to them indicate the vibration frequency. Results of linear regression analysis are summarized in Table 2.

c) Integrated SM EMG over the interval 120-250 ms.

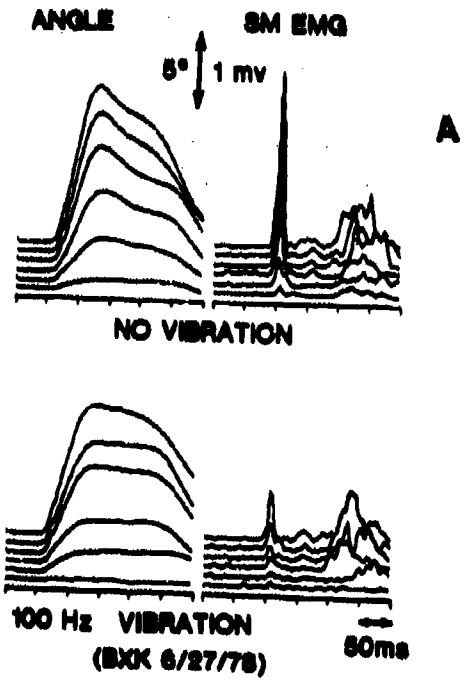
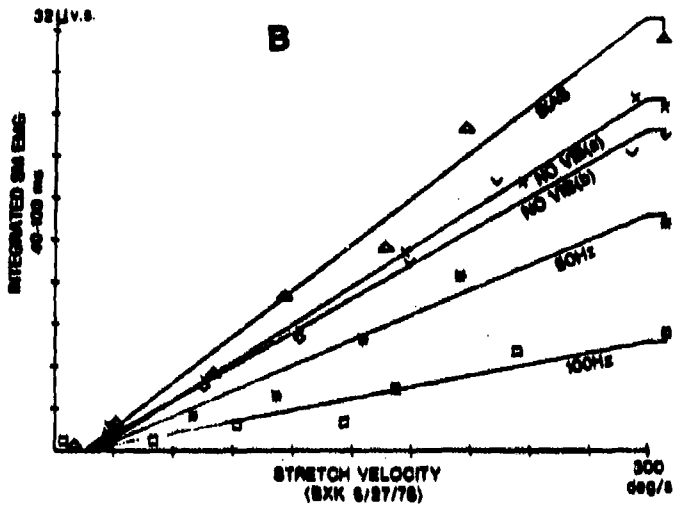


Fig. 8: a) Angular displacement and SM EMG for another subject. The lower group is under the condition of 100 Hz vibration.



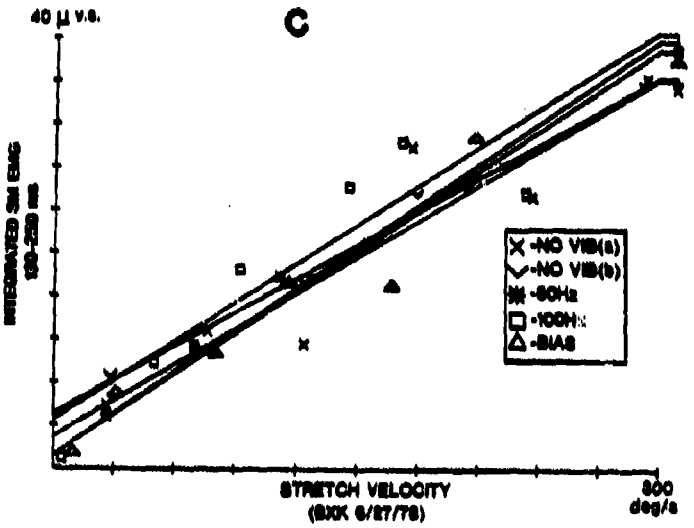


Fig. 8: b) Integrated SM EMG over the interval of 40-100 ms. The five data sets were no vibration (a) at the start of the experiment, vibration at 50 Hz and 100 Hz with TVR counterbalanced, voluntary bias without vibration, and no vibration (b) at the end of the experiment. Results of linear regression analysis are summarized in Table 2.

c) Integrated SM EMG over the interval 130-250 ms.

INPUT ADAPTIVE BEHAVIOR OF THE HUMAN EYE TRACKING SYSTEM

Loek Kratzschmar
Department of Physiology
Faculty of Medicine
Erasmus University Rotterdam
3000 DR Rotterdam, The Netherlands

The frequency response of the human eye tracking system appears to be highly dependent on the ability of the subject to predict the future course of the target. Bode plots obtained from experiments with simple sinusoids show a much smaller phase lag (sometimes even a small phase lead) and higher gain than those where more demanding signals such as band limited noise or a sum of sinusoids have been used. According to some investigations this phenomenon, which has also been observed in other motor tasks like manual control, vergence eye movements and accommodation, is a continuous function of the target motion "complexity" and is often attributed to a predictor mechanism in the human brain. In the present paper the exact nature of this "predictive" or "precognitive" behavior has been studied. It was found that this behavior is merely a matter of an adaptive proces and is not related to subjective predictability of the target trajectory. The tracking system can be described as an adaptive servosystem of which the various parameters and system type are adjusted according to some optimal proces to obtain a tracking error and error rate which are as small as possible. The time delay of the system however, has a rather constant value for all continuous input signals and is only affected by fatigue and blurring of the retinal image.

ELECTRODERMAL LABILITY AND CAPACITY FOR DUAL-TASK PERFORMANCE

by
Russell A. Benel
Michael G. H. Coles
Denise C. R. Benel
Department of Psychology
University of Illinois at Urbana-Champaign
Champaign, IL 61820

ABSTRACT

Electrodermal (galvanic skin) responses have long been used as a convenient index of sympathetic arousal. Individuals exhibiting high levels of resting electrodermal activity are designated "labiles," while those with low levels, "stabiles." Labiles have been found to be resistant to performance decrement over time in vigilance studies. Thus, increased electrodermal responsivity appears to represent enhanced attentional and information processing capacity during such tasks. However, it is not clear whether there is a cost attendant on this enhancement or whether the performance of labiles is enhanced on other tasks. The responses of 10 labiles and 10 stabiles were collected during a simple reaction time task and under single and dual-task monitoring and tracking conditions. The monitoring task involved the detection of dynamic system failures in an automatic tracking task. The tracking task was the Critical Task with a sub-critical level of instability. The reaction time data parallel the previously reported findings of inferior performance for stabiles. The data for complex single-task and dual-task conditions did not reveal a similar trend. Generally, labiles performed better in both single and dual conditions. The pattern of these results suggest the proposed selective enhancement associated with electrodermal lability is only adaptive for tasks that require or profit from focused attention. The implication for monitoring tasks is discussed.

INTRODUCTION

Much previous research has been devoted to the study of man as the physical manipulator at the controls of complex man-machine systems. A consistent and often not so subtle trend has been the shift in the role of man from that of manual controller toward man as supervisor (1) or monitor of automatically controlled functions. Removal of man from an active element in the control loop has led to questions concerned with man's ability to monitor such systems for failure and to regain control of such systems when failed. Whether man is actively controlling or not, the essential task remains of monitoring all facets of system performance.

The existence of individual differences in either manipulating or monitoring performance is not surprising. For many investigators with interests in man-machine systems, this intersubject variability is more frequently a source of consternation rather than a topic for research. Additionally, investigators with primary interests in individual differences are not often concerned with man-machine systems. Nevertheless, there is one area in which some overlap may be seen. Previous research within the vigilance and signal detection paradigms has shown that there is a general decline in performance over time known as vigilance decrement. It has also been found that some subjects do not exhibit this decline. The parallel between vigilance (and/or signal detection) and monitoring may be drawn without undue difficulty.

The quest for an individual characteristic that could be consistently related to monitoring performance has led to an interest in electrodermal (galvanic skin) responses. As Katkin (2) has succinctly stated, the logic for the investigation of the relationship between electrodermal lability and attentional or vigilant behavior lies in the consistent positive relationship between the rate of spontaneous electrodermal responses (electrodermal lability) and the rate of habituation of the attention-related orienting response. Crider and Augenbaum (3) suggested that the individual differences in electrodermal habituation (operationally equivalent to lability) can be considered to reflect characteristic rates of attentional decrement with repetition of the eliciting event. Therefore, fast habituators (low lability or stabiles) should show a performance decrement over time not observed among slow habituators (labiles) in tasks requiring sustained attention.

Evidence is available to support this hypothesis. Coles and Gale (4) found a positive correlation between lability and auditory signal detection performance. Siddle (5) reported a difference between stabiles and labiles in the time course of auditory vigilance performance. Both groups detected an equal number of signals early but stabiles exhibited a decrement over time not seen among labiles. Crider and Augenbaum noted an ambiguity in the relationship between lability and attentional capacity. They argued that differences in performance could be due to differences in perceptual sensitivity (d') or response bias (β). Although Crider and Augenbaum replicated the previous finding of labiles detecting more signals, they attributed it to

an increased willingness to respond. That is, increased correct detections were accompanied by increased false alarms. Their initial data tend to support the finding that d' remains constant across time for labiles, but declines for stabiles. The replication and modification of procedures suggested that motivational factors expressed themselves through a relatively liberal reporting strategy.

Katkin and his colleagues reported the results of a similar experiment in which the same auditory signal detection task used by previous researchers was performed under differential payoff schedules (6). They replicated the finding of performance differences between labiles and stabiles, but found no difference in response bias. Furthermore, the payoff matrix did not have differential effects on detections or d'. When the data were analysed by blocks of time it was found that hits and d' declined across successive blocks for stabiles but not for labiles. The greater apparent internal consistency of Katkin's data suggests that some attentional factor rather than a bias factor is responsible for the performance differences.

Although labiles appear to have superior vigilance performance at least in single task conditions, it is possible that this performance is gained and maintained at the expense of other performance. Katkin proposed that his data suggest that lability reflects a variable that is a selective enhancer of effective central processes and not generalised arousal. Therefore, increased lability may reflect enhanced attentional or information processing capacity. However, this might be quite specific in the manner of selective attention maintenance. In simple single task conditions performance may be enhanced, whereas in dual tasks (or more complex single tasks) such enhancement may come at the expense of other aspects of the tasks. When multiple performance measures are available, it may be possible to determine if the hypothesised selective enhancement has any attendant costs. Or alternatively, whether such selective enhancement extends beyond the realm of vigilance.

METHOD

To assess the relationship between electrodermal lability and performance an experiment was conducted in which subjects who differed in lability performed in a paradigm that allowed collection of performance measures similar to those found in the vigilance paradigm as well as other performance measures.

Subject Selection

An initial unselected sample of 59 male undergraduates was randomly drawn from the University of Illinois, Department of Psychology subject pool. The only restriction was that all would be right-handed. Each of these subjects

was screened for electrodermal lability under standardized conditions. An initial 5 min resting measure of lability was taken, followed by a measure during a 5 min period in which subjects performed a simple reaction time (RT) task (key press to light stimulus), ending with a 5 min post-task resting measure. During the entire procedure the subject was seated in a lounge chair in a semi-reclined position placed in a sound attenuated chamber. Subjects were cautioned to remain quiet and still during the resting periods. For the task period the chamber lights were turned off following instruction in the RT task.

Two subgroups of subjects were selected on the basis of the number of spontaneous fluctuations in skin conductance of at least 0.5 μmho during the rest period. The ten highest and lowest responders were chosen to represent labiles and stabiles respectively. In Figure 1, the responses of typical labile and stabile subjects are shown for the three measurement periods. Selected Ss proceeded to a second stage of the experiment in which they performed tracking and monitoring tasks.

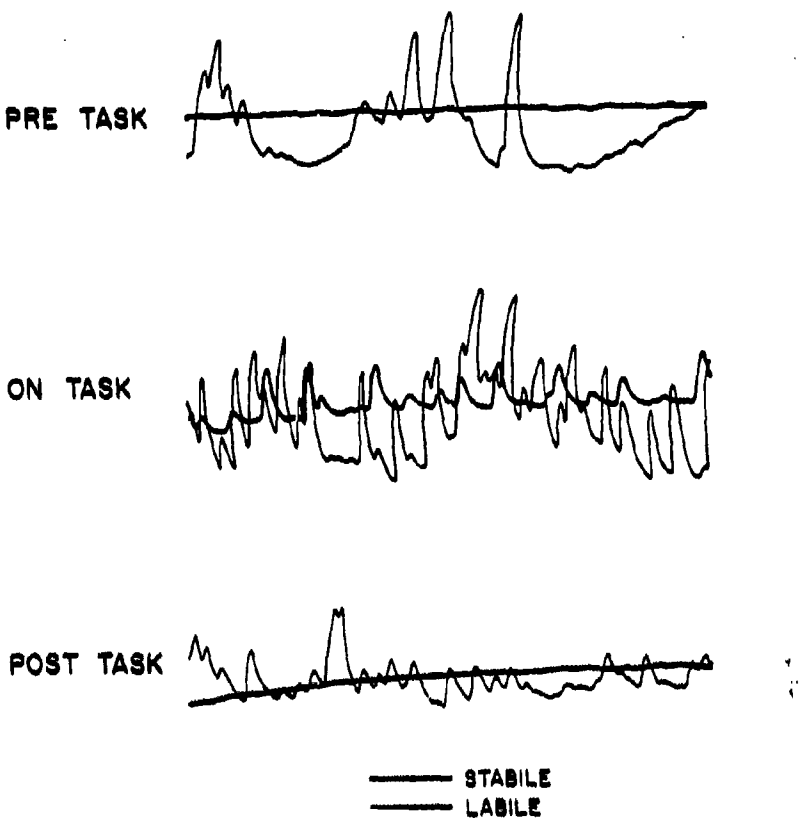


FIGURE 1. Typical electrodermal activity of labile and stabile subjects during resting and reaction time task conditions.

Apparatus and Procedure

Measures of skin conductance were obtained using Beckman electrodes attached to the volar surface of the middle and index fingers of the left hand. A 9852 A Beckman coupler which imposes a constant voltage of .5v provided a tracing of the subject's conductance (sensitivity: 1mm = .1 μmho).

The reaction time task was programmed on a DEC LSI-11 microprocessor. Stimuli were presented with a random interstimulus interval of 4 to 5 sec and subjects were allowed 1 sec to respond. The subject held the stimulus/response box in his right hand, using his thumb to respond on the push button immediately below the stimulus light. A task light stayed on continuously during the task. Each subject was given fifteen practice trials prior to the 50 experimental trials. All data was stored during and printed after the task by the microprocessor.

The dual task performance apparatus is described in detail elsewhere (7). It consisted of a 7.5 x 10 cm Hewlett-Packard Model 1300 CRT display, a spring-centered, dual axis joystick with trigger, and a spring-loaded finger controller. A Raytheon 704 digital computer generated tasks and processed responses. Subjects sat in an arm chair in a sound attenuated chamber with the dual axis joystick on the right and the finger controller on the left. The display subtended 3.4° of visual angle. The two basic tasks for the subject were the Critical Task controlled by the left hand and the detection of failure in an automatic tracking task signalled by trigger presses. The automatic tracking task was a pursuit task in which the target followed a semi-predictable path. The computer's tracking behavior modelled that of a human controller. For the conditions described below failures were step changes in the acceleration constant (α) toward a system approximating pure second order dynamics. For the Critical Task the value of the instability parameter was set at a constant subcritical level ($\lambda = 1.5$) for all conditions analyzed to date.

The subjects in each of the subgroups participated in three sessions. The first was a training session followed on two separate days by two experimental sessions. On each test day there were three practice and 16 experimental trials. The conditions which have been analyzed include two trials of the Critical Task (CT) alone, two trials of automatic monitoring (AU) alone for failure in detection, and 4 trials of dual tasks (AU + CT).

Failures (6-8 per 2.5 min trial) were generated by an algorithm assuring random interfailure intervals and low system error. The latter precluded obvious "jumps" in cursor position. Detections were recorded by trigger presses which reset the dynamics. If no response occurred within 6 sec the system returned to normal via a 4 sec ramp. Detections were assessed in terms of accuracy and latency. Hit responses were those detection responses in the 6 sec interval (P(Hit) = Hits/intervals). P(FA) was the number of false alarms/false alarm intervals. Rather than d' the nonparametric area under the ROC curve (P(A)) was used as the sensitivity measure (8). Latencies to detection were also recorded. Average absolute error was recorded for the Critical Task.

For the reaction time task performed by the initial unselected sample the correlation between RT and electrodermal lability was statistically reliable (r (59) = -.34; p < .01). This suggests that the previously documented relationship between vigilance and lability exists for RT. When the labile and stabile subgroups are compared for the RT performance across time, Figure 2 indicates a decrement across time for stabiles, which is greater than that for the labiles. However, while the effects of both lability (p < .01) and trials (p < .003) are reliable, the interaction was not (p > .05).

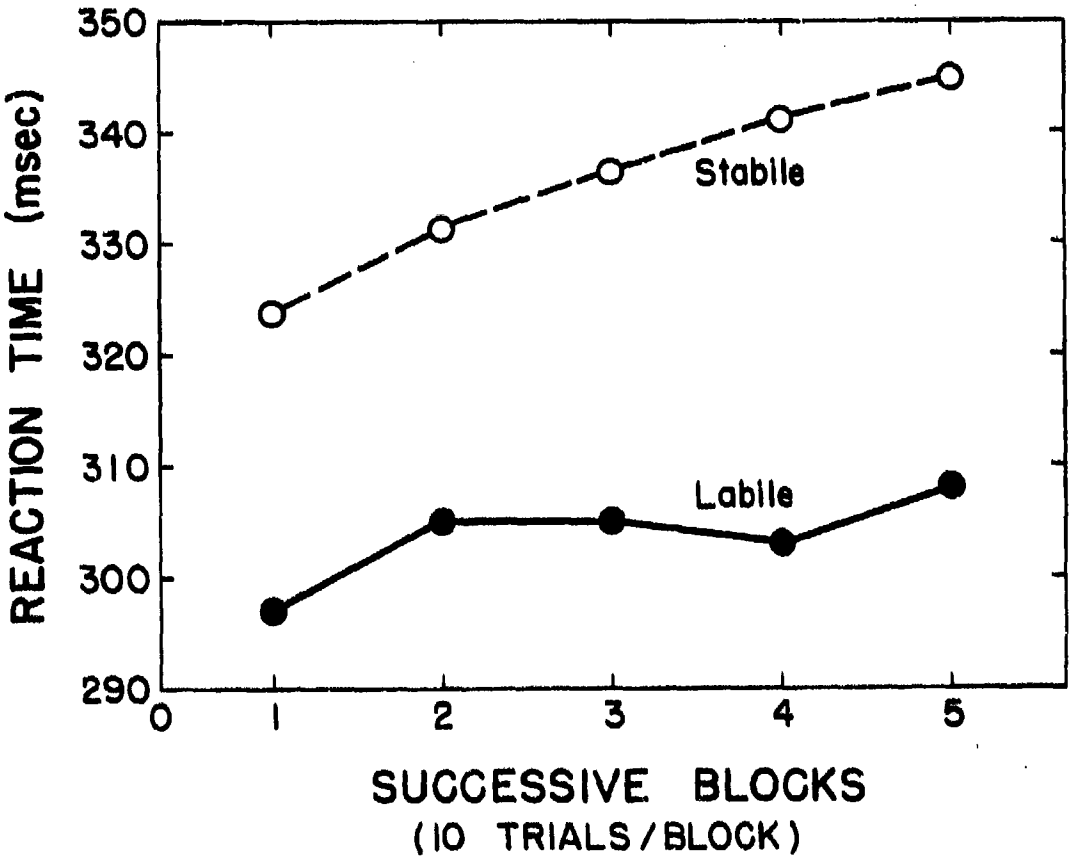


Figure 2. Reaction times for labile and stabile subjects as a function of time on task.

Under dual task conditions the relationship is less clear. In Figure 3 the detection performance measure of sensitivity P(A) (≈ d' in signal detection) is plotted against the average absolute error on the Critical Task (CT). Both groups show a decrement under dual vs single task conditions resulting primarily from increased CT error. This decrement is highly reliable ($p < .00001$). However, there was no reliable difference between labiles and stabiles although stabiles tended to perform better ($p = .18$).

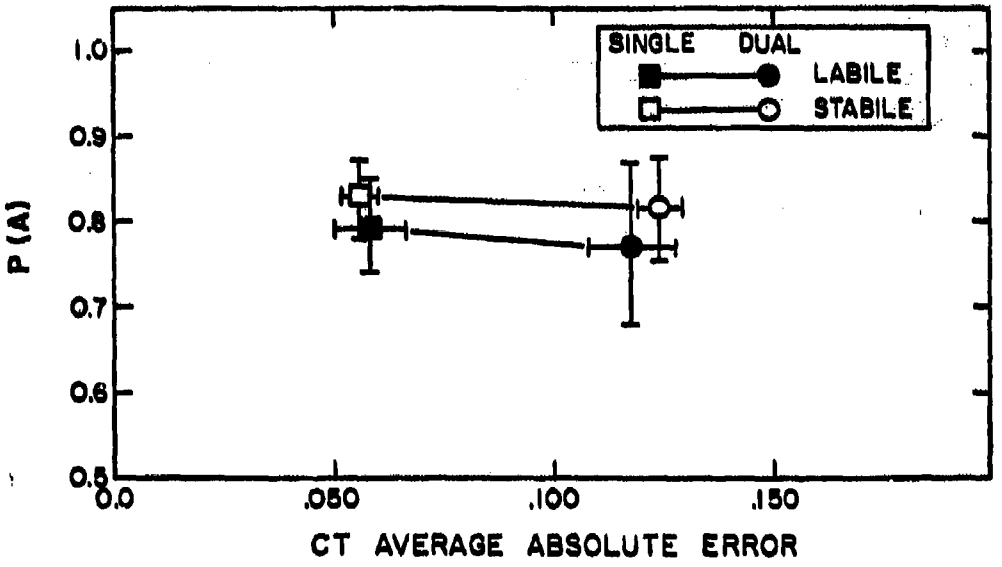


Figure 3. Sensitivity (P(A)) and average absolute error value for labile and stabile subjects under single and dual task conditions. Standard deviation values are shown for each point.

When the two measures for the detection task, P(A) and latency, are compared (see Figure 4), stabiles also appear to show superior performance. However, the main effect for lability ($p = .18$) as well as the effects for single vs dual task ($p = .14$) and their interaction ($p = .10$) were not significant. A Fisher Exact Probability Test on the frequency of subjects within each subgroup showing a dual task decrement indicates a tendency for labiles to be insensitive to the dual task manipulation ($.1 > p > .05$).

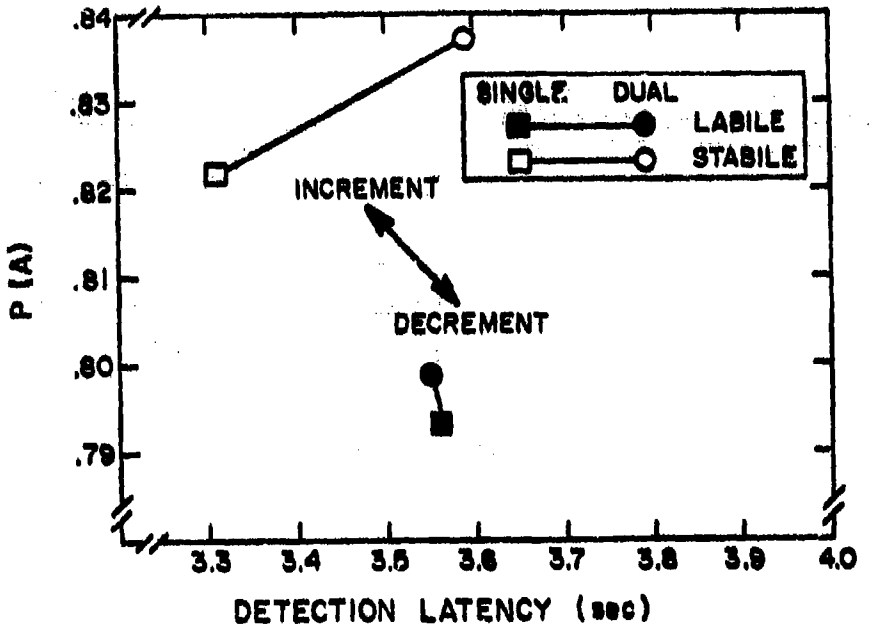


Figure 4. Sensitivity (P(A)) and detection latency values for labile and stabile subjects under single and dual task conditions. The direction of increment and decrement in performance is shown on the negative diagonal.

DISCUSSION

The pattern of data suggest that Katkin (2) may be correct in assuming the electrodermal lability reflects a variable that is a selective enhancer of effective central processes and not generalized arousal. Certainly vigilance (e.g., 4, 5, 6) and reaction time tasks have been shown to be performed better by labiles. However, the mechanism appears to be related to focused attention. Tasks which require or allow focused attention show a benefit associated with lability. On the other hand, under increasingly complex single task and/or dual task conditions focused attention appears to be a maladaptive rather than an effective central process.

The complexity of the monitoring task used in this study affords ample opportunity to focus on irrelevant dimensions. Wickens and Kessel (7) have argued that this detection task requires primarily perceptual and decision making components and any misdirection of these resources should be manifested

in performance decrements. A focused attention response under conditions that demand flexibility in attention switching or division would lead rapidly to performance deterioration. In fact, labiles' performance is not only somewhat poorer it is also somewhat more variable. This is most plausibly explained in terms of only occasional focusing on task relevant dimensions.

One obvious difference between the current monitoring task and the previously used vigilance tasks is the reduction in signal presentation for the current task (down to six to eight presentations per 2.5 min trial). The reaction time task whose results paralleled the previous vigilance findings presented signals at a rate more nearly equal to the vigilance task. The role of signal rate in vigilance tasks has been investigated, but not in relation to electrodermal activity. Other differences include the continuous nature of the monitoring task versus the discrete nature of reaction time or the vigilance task (albeit with a memory component), blocks of trials versus uninterrupted event periods, and signal discriminability.

Although the data of the present experiment must be regarded as preliminary, they do indicate the possible value of the measurement of electrodermal lability as a method of identifying individual differences in attention related tasks. These data suggest that labiles are superior on focused attention tasks but their performance suffers when flexibility is required. Future research should examine the relationship between lability and the dimension of allocation flexibility. It would be expected that the superior ability of labiles to focus attention would be an advantage under some circumstances, but a serious disadvantage under others. Such research would also indicate the utility of the electrodermal lability in the selection of human operators.

REFERENCES

1. Sheridan, T. B. Preview of models of the human/monitor/supervisor. In T. B. Sheridan and G. Johannson, (Eds.), Monitoring behavior and supervisory control. New York: Plenum, 1976.

2. Katkin, E. S. Electrodermal lability: A psychophysiological analysis of individual differences in response to stress. In L. G. Sarason and C. D. Spielberger, (Eds.), Stress and anxiety, Vol. II. Washington, D.C.: Hemisphere, 1975.

3. Crider, A. and Augenbaum, C. D. Auditory vigilance correlates of electrodermal response habituation speed. Psychophysiology, 1975, 12, 36-40.

4. Coles, M. G. H. and Gale, A. Physiological reactivity as a predictor of performance in a vigilance task. Psychophysiology, 1971, 8, 594-599.

5. Siddle, D. A. T. Vigilance decrement and speed of habituation of the GSR component of the orienting response. British Journal of Psychology, 1972, 63, 191-194.

6. Sostek, A. J., Katkin, E. S., and Sostek, A. M. Signal detection as a function of electrodermal lability and differential payoffs. Paper presented at meeting of the Society for Psychophysiological Research, Boston, November, 1972.

7. Wickens, C. D. and Kessel, C. The effects of participatory mode and task workload on the detection of dynamic system failures. Savoy, IL: University of Illinois at Urbana-Champaign, Aviation Research Laboratory, Technical Report ARL-77-8/AFOSR-77-7, June, 1977.

8. McNicol, D. A primer of signal detection theory. London: Allen and Irwin, 1972.

9. Wickens, C. D. and Kessel, C. The processing resource demands of failure detection in dynamic systems. Urbana, IL: University of Illinois at Urbana-Champaign, Engineering-Psychology Research Laboratory, Technical Report, EPL-79-1/AFOSR-79-1, January, 1979.

10. Kahneman, D. Attention and effort. Englewood Cliffs, N.J.: Prentice Hall, 1973.

PREDICTOR OPERATOR IN PURSUIT AND COMPENSATORY TRACKING

By

Robert J. Jaeger
Gyan C. Agarwal
Gerald L. Gottlieb
College of Engineering
University of Illinois at Chicago Circle
Chicago, Illinois 60680
and
Department of Physiology
Rush Presbyterian-St. Lukes Medical Center
Chicago, Illinois 60612

ABSTRACT

The predictor operator is a well established property of a human controller in manual control systems with a pursuit display. For periodic inputs, the human controller has been shown to use a synchronous tracking mode with considerable system performance improvement. The performance change is most evident in the dramatic change in the phase portion of the system's bode plot. In the present study, we investigated the predictor operator in a manual control system with a compensatory display and compared it to similar results with a pursuit display. Four types of input signals were used: 1) single sinusoid, 2) narrow band summed sinusoids, 3) variable band summed sinusiods, and 4) wide band summed sinusoids. The frequency range was from 0.04 to 1.5 Hz. Our data from pursuit tracking is similar to data obtained in several previous studies. What is interesting is that the predictor operator in compensatory tracking is essentially the same as in pursuit tracking. The predictor operator is dependent on the input bandwidth.

INTRODUCTION

The predictor operator is a well established property of the human controller in manual control systems with a pursuit type of display. If the system input is periodic, this periodicity may be detected, and the controller uses the predictor operator to enter what has been termed the synchronous tracking mode. In the synchronous tracking mode, performance is improved considerably over the case in which system input is not predictable.

This performance change is most evident in the dramatic changes in the phase portion of the system's bode plot. Examples of this have been published for a two pointer manual tracking task and for smooth pursuit eye movements in response to a smoothly moving visual target (1,2). In each of these examples, a marked improvement was seen between phase curves obtained when the input was predictable. The difference in phase curves has been attributed to the action of a predictor operator.

In the present study, we wished to investigate the predictor operator in a manual control system with a compensatory display, and make comparisons to similar investigations with a pursuit display. The existence of a predictor operator in compensatory tracking has not been investigated. Most investigators seem to prefer a pursuit display when studying the predictor operator. Our hypothesis is that a predictor operator is also present in manual control systems with a compensatory display.

METHODS

Four subjects were used in the present study. Prior to collection of data, subjects were run through a scaled down set of experimental blocks to permit practice and allow for the approach of asymptotic performance levels.

The experimental setup is shown schematically in figure 1. The human operator sat before an oscilloscope and manipulated a one degree of freedom joystick. The joystick rotated freely with no springs or dashpots. The plant dynamics consisted of a simple unity gain. Two horizontal lines were displayed on the oscilloscope, and were driven vertically.

The first line served as the target in the pursuit case, and a zero error reference line in the compensatory case. The other line displayed system output in the pursuit case, and system error in the compensatory case. The lines were adjusted in focus and intensity which made it easy for subjects to discriminate between the two. The display configuration could be changed between pursuit and compensatory by means of a switch. Error was computed by a simple analog circuit. The entire experiment was under the control of a real time digital computer (General Automation SPC-16/65) with digital-to-analog and analog-to-digital channels. The computer operated at a 20 msec sampling time, generating the appropriate system input and monitoring the system output. There was a 2.54 second delay between the initiation of the input signal and the recording of the output signal to exclude the start-up transient. A total of 61.44 seconds of data were taken for each tracking period. Phase angle between system input and system output as a function of input frequency was computed by standard fourier series techniques. The root mean square (RMS) value of the system error was digitally computed using the equation:

$$\text{RMS Error} = \frac{1}{N} \sum_{k=1}^{N} (y_k - x_k)^2$$

where:

y_k = kth output sample

x_k = kth input sample

N = Total number of samples (3072)

The RMS of the remnant was digitally computed using the following equation:

$$\text{RMS Remnant} = \frac{1}{N} \sum_{k=1}^{N} \left[y_k - \sum_{j=1}^{p} g_j a_j \sin(\omega_j k \Delta t + \theta_j) \right]^2$$

where: p= number of sinusoids in the input signal

a_j, ω_j= amplitude and radian frequency of the jth input sinusoid

g_j, θ_j= gain and phase angle of the output at the input frequency ω_j

Δt = sampling interval (0.02 second)

Note that the remnant was assumed to be a noise signal added to the human operator's output, as is typically done in modeling of this nature.

Four different classes of system inputs were used with each display configuration. The frequency range of interest was 0.04 to 1.6 Hz. The four classes of inputs were (I) single sinusoids, predictable; (II) narrow band sum of sinusoids, quasi-predictable (III) variable band sum of sinusoids, quasi-predictable; and (IV) wide band sum of sinusoids, unpredictable. Examples of these inputs are shown in figure 2. For predictable input I, phase lags and the two RMS performance measures were determined for at least three replications of seven input frequencies. The results were averaged at each frequency. For quasi-predictable input II, the predictability of the single sinusoid was degraded by adding one higher and one lower frequency component to the test frequency (3). Phase lag measurements were made only at the center frequency. The two RMS performance measures were made over each tracking period and were considered measurements at the center frequency. Five different center frequencies were used with a minimum of three replications at each center frequency. For quasi-predictable input III, predictability was degraded by adding to the test frequency between one to four sinusoids at lower frequencies. Phase lag measurments were made only at the test frequency. Three or more replications were made for each of five variable band inputs. For unpredictable input IV, five frequencies were used by forming several different combinations out of a set of seven frequencies over the frequency range. A minimum of three replications were run for each combination. Phase lag measurements were made at each frequency in a given

tracking period. Error and remnant RMS performance measures were made for each tracking period.

The sum of amplitudes of individual sinusoids in the input signal was always equal to ten. When more than one sinusoid was summed in the input signal , the component sinusoids were adjusted to be equal in amplitude.

The pursuit data were taken first with the order of the four classes of inputs randomized between subjects. Compensatory data were then taken, again with the four classes of inputs randomized for each subject.

RESULTS

The phase portion of the system's bode plot from input to output was constructed for all data. Phase lag in degrees on a linear scale was plotted against frequency on a log scale. Two bode plots were prepared for each subject--one for the pursuit display and one for the compensatory display. Four phase curves appear on each plot corresponding to the four classes of input used. The circle is for input I, predictable sinusoids, and the square is for input IV, unpredictable sums of sinusoids. The two quasi-predictable inputs II and III are represented by diamonds and triangles respectively. The bode plots are given in figures 3 to 6.

Normalized remnant was plotted on a linear scale against frequency on an arbitrary scale. The plot symbols are identical to those used on the bode plots, except that the average RMS across all wide band tracking periods is shown as a dashed line. Again, two plots were made for each subject, one for pursuit and one for compensatory. The remnant plots are given in figures 7 to 10. The normalized RMS error plots were prepared in a manner similar to the remnant plots and are given in figures 11 to 14.

DISCUSSION

There are two other types of predictor operators that have received attention in manual control. Precognitive manual control displays differ from the pursuit and compensatory displays of the present study in that system input may be previewed by the subject. This allows the operator to preprogram his responses (4). A second type of predictor operator has been investigated in control systems with a periodic input. After tracking has been established, the input is turned off. The controller has been instructed to maintain his response as if the input were still present (5). These two types of predictor operators may ultimately be related to the predictor operator of the present study. However, in terms of analysis, each of these predictor operators must be analyzed somewhat differently.

Our pursuit phase curve results are very similar to previous investigations of this natue (2,6). All of our subjects exhibited a similar pattern of pursuit phase curves. The single sinusoid phase curve was typically flat around zero degrees across the frequency range of interest. The wide band phase curve fell off rapidly with frequency. The two quasi-predictable phase curves lay between the predictable and unpredictable extremes. Some occasional overlap was observed at low and high frequencies betweeen the quasi-predictable phase curves and the single sinusoid or wide band phase curves. In no case, however, did single sinusoid and wide band phase curves overlap.

The compensatory phase curves were somewhat similar to the pursuit phase curves. The single sinusoid phase curve fell off gently with frequency and the wide band phase curve fell off rapidly with frequency. The quasi-predictable phase curves lay between the two.

When comparing pursuit phase curves with their compensatory counterparts, the pursuit curves represent a generally better level of tracking performance. That is, there is less phase lag in the pursuit case. Similar differences between performance measurements of pursuit and compensatory displays have been observed (7). For simple gain dynamics, the pursuit display seems to be preferred over the compensatory display, although, for higher order plant dynamics the compensatory display may be better (8).

The separation of phase curves in the pursuit case as a function of input predictability shows the action of the predictor operator, as mentioned in the introduction. Note that a similar separation is observed between the compensatory phase curves, also indicating a predictor operator.

In contrast to the phase curves however, the normalized RMS remnant and normalized RMS error plots do not show the dramatic and consistent separation with graduation in predictability as the phase curves do. Note that curves for single sinusoid, narrow band and variable band inputs all approach and in many cases exceed the wide band values.

It has been suggested (7) that the remnant signal should be small when the controller is successfully predicting and large when prediction is poor. Our results are consistent with this suggestion only at the lower frequencies of our study. At mid and higher frequencies, normalized RMS remnant values for predictable and quasi-predictable inputs approach and typically exceed the wide band RMS remnant value.

The normalized RMS error results are similar to the results for normalized for RMS remnant. Several authors (7) have suggested that the predictor operator may be demonstrated by observing reduced error scores of some type. Again, our data are

consistent with this suggestion only at low frequencies.

These seeming inconsistencies may be resolved if one considers an hypothesis that the human controller first concentrates on making responses in the correct direction. When this has been achieved, the operator next attends to matching the timing of the input signal and response. Finally, attention is turned to achieving accurate responses in amplitude (8,9,10). The first two of these three criteria contribute primarily to the phase measurement. The data of the present study support the contention that the phase curve is the best indicator of the predictor operator.

Returning to the pursuit phase curves, consider the difference between the wide band and single sinusoid pursuit phase curves. This difference is generally agreed to be the result of the predictor operator.

Notice that a similar difference exists between the wide band and single sinusoid compensatory phase curves even though the compensatory phase curves are below their pursuit counterparts. It is the separation of the wide band and single sinusoid phase curves that is similar between the pursuit and compensatory case. This difference was defined as phase improvement and plotted versus frequency for both pursuit and compensatory displays in figure 15. What is interesting is that the phase improvement curve for each display configuration is similar. This could suggest that the same mechanism was responsible in each case for the phase improvement. In the pursuit case, periodicity of the input is presented to the subject visually as well as proprioceptively and kinesthetically from the stick movement. In the compensatory case, periodicity of the input signal is presented to the subject proprioceptively and kinesthetically from the stick movement, since no periodic visual display is present. Despite this loss of visual information in the compensatory case, the predictor operator provides phase improvement to a similar degree in both pursuit and compensatory tasks.

We conclude that the predictor operator improves the system's phase response by comparable amounts in both the pursuit and compensatory case, but is much less effective in improving the two RMS performance measures.

ACKNOWLEDGEMENT

This work was supported by NSF Grant ENG-7608754 and NIH grants NS-00196 and NS-12877.

REFERENCES

1. Stark, L. Neurological Control Systems, New York, Plenum Press, pp 247 and 334, 1968.

2. Dallos, P.J. and Jones, R.W. Learning behavior of eye fixation control system. Institude of Electrical and Electronics Engineers Transactions on Automatic Control, vol 8, pp 218-227, 1963.

3. Magdeleno, R.E., Jex H.R., and Johnson W.A. Tracking Quasi-Predictable Displays: Subjective Predictability Gradations, Pilot Models for Periodic and Narrow Band inputs. Proceedings of the Fifth Annual Conference on Manual Control, National Aeronautics and Space Administration, NASA SP-215, pp 391-428, 1970.

4. Poulton, E.C. Learninng the statistical properties of the input in pursuit tracking. Journal of Experimental Psychology, vol 54, pp 28-32, 1957.

5. Vossius G. Der Kybernetische Aspekt der Willkurbewegung, Progress in Cybernetics, New York, Elsevier, 1975.

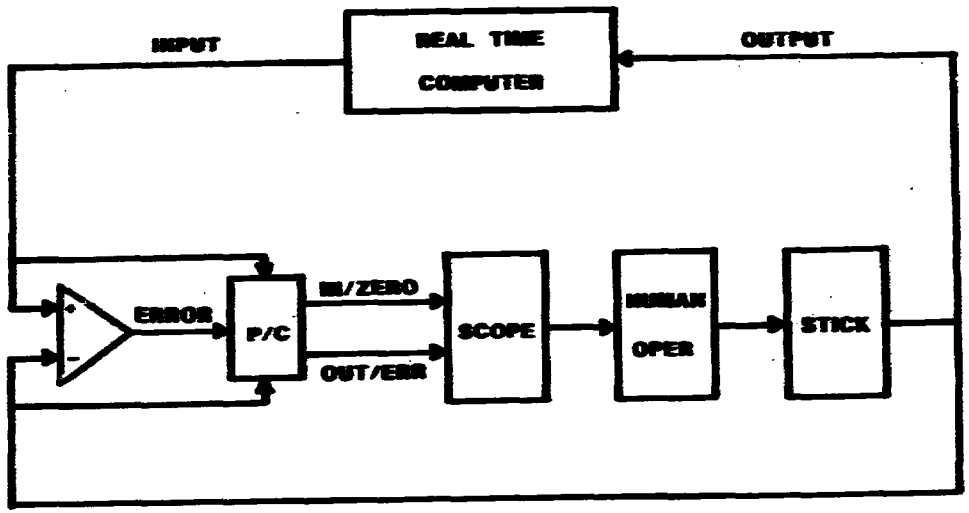
6. Elkind, J.I. Characteristics of simple manual control systems. Massachusetts Institute of Technology, Lincoln Laboratory, Technical Report 111, 1956.

7. Poulton, E.C. Tracking skill and manual control, New York, Academic Press, chapters 8-9, 1974.

8. Allen, R.W. and Jex H.R. An experimental investigation of compensatory and pursuit tracking displays with rate and acceleration control dynamics and a disturbance input. National Aeronautics and Space Administration, Contractor Report NASA CR-1082, 1968.

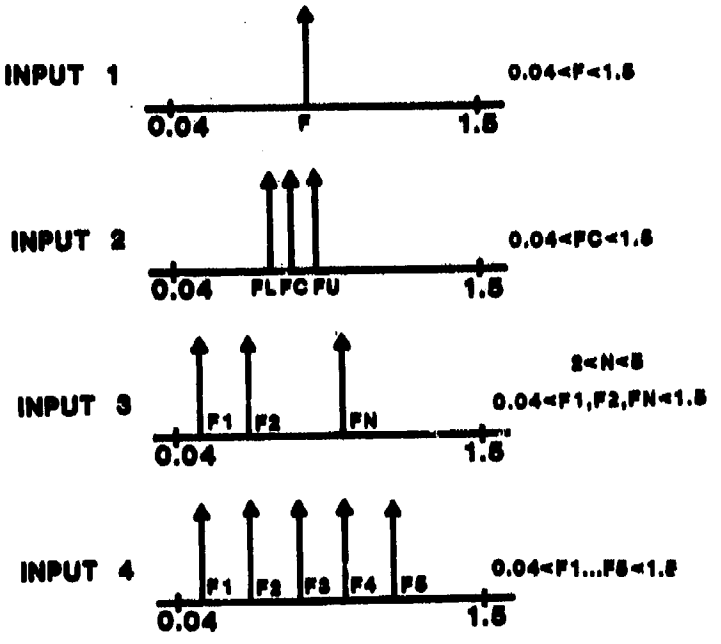
9. Noble, M., Trumbo, D., Ulrich, L., and Cross, K. Task predictability and the devvelopment of tracking skill under extended practice, Journal of Experimental Psychology, vol 72, pp 85-94, 1966.

10. Noble, M. and Trumbo, D. The organization of skilled response. Organizational Behavior and Human Performance, vol 2, pp 1-25, 1967.



SCHEMATIC OF EXPERIMENT

Figure 1



SCHEMATIC OF INPUT

Figure 2

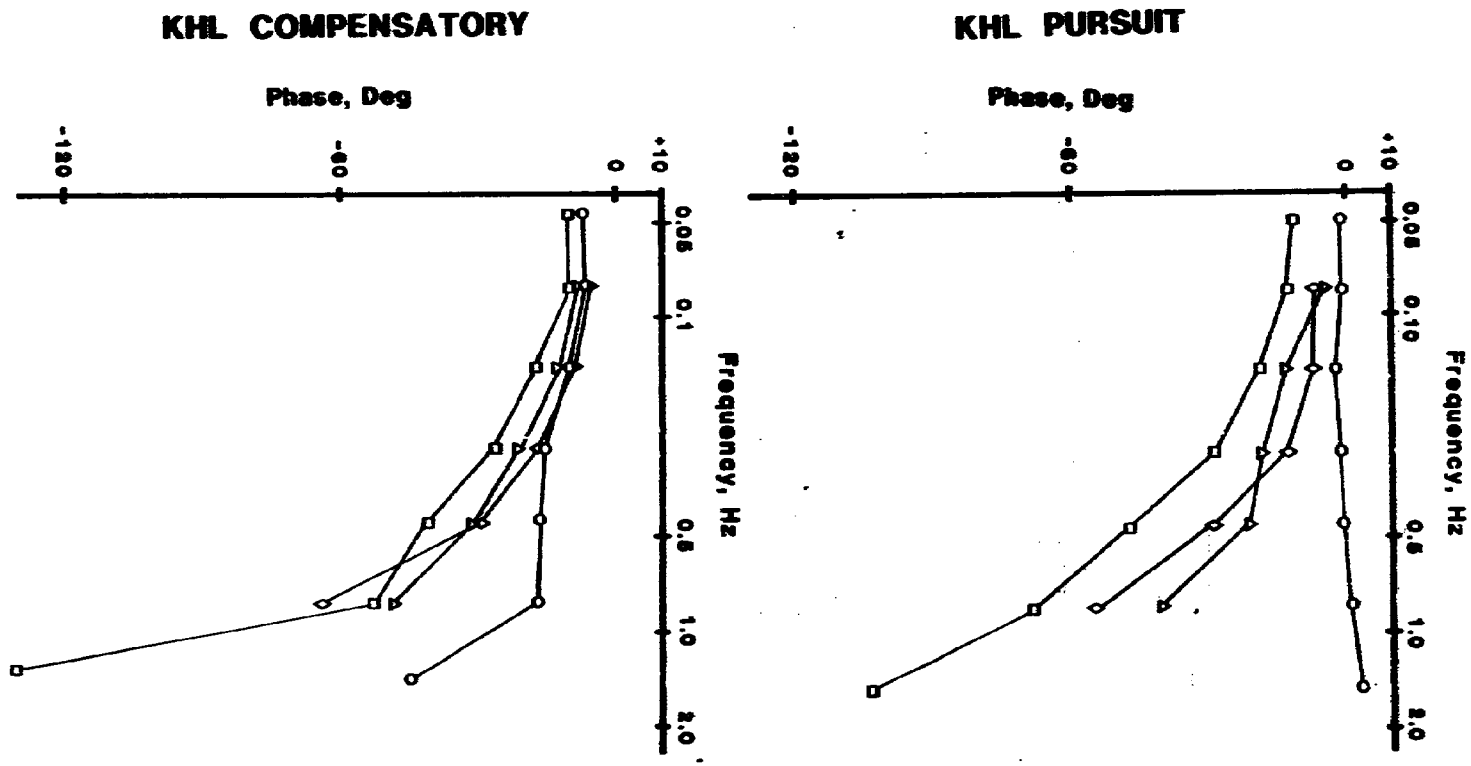

KHL COMPENSATORY
Phase, Deg
-120
-60
0
+10
0.05
0.1
0.5
1.0
2.0
Frequency, Hz
KHL PURSUIT
Phase, Deg
-120
-60
0
+10
0.05
0.10
0.5
1.0
2.0
Frequency, Hz

Figure 3

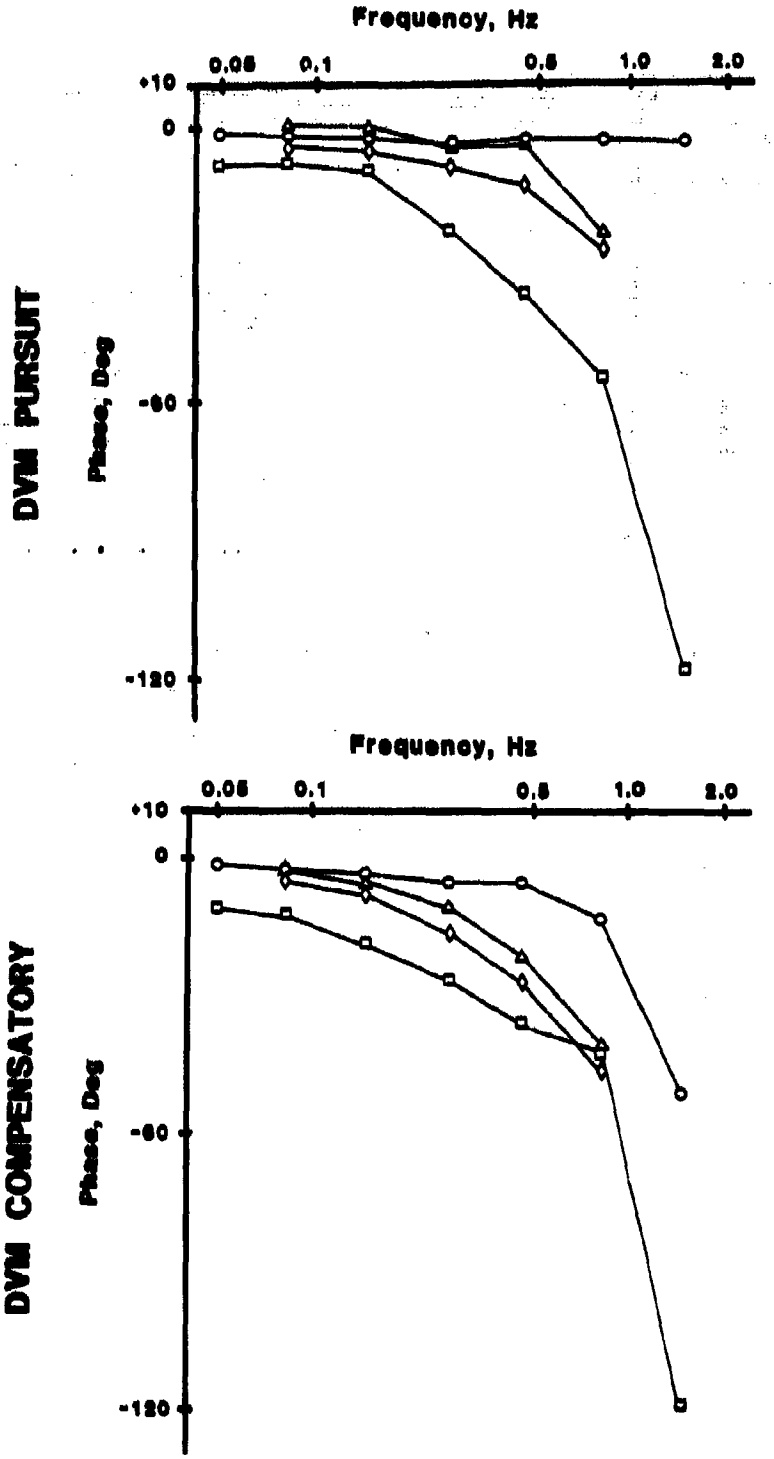

DVM PURSUIT
Phase, Deg
Frequency, Hz
0.05
0.1
0.5
1.0
2.0
+10
0
-60
-120
DVM COMPENSATORY
Phase, Deg
Frequency, Hz
0.05
0.1
0.5
1.0
2.0
+10
0
-60
-120

Figure 4

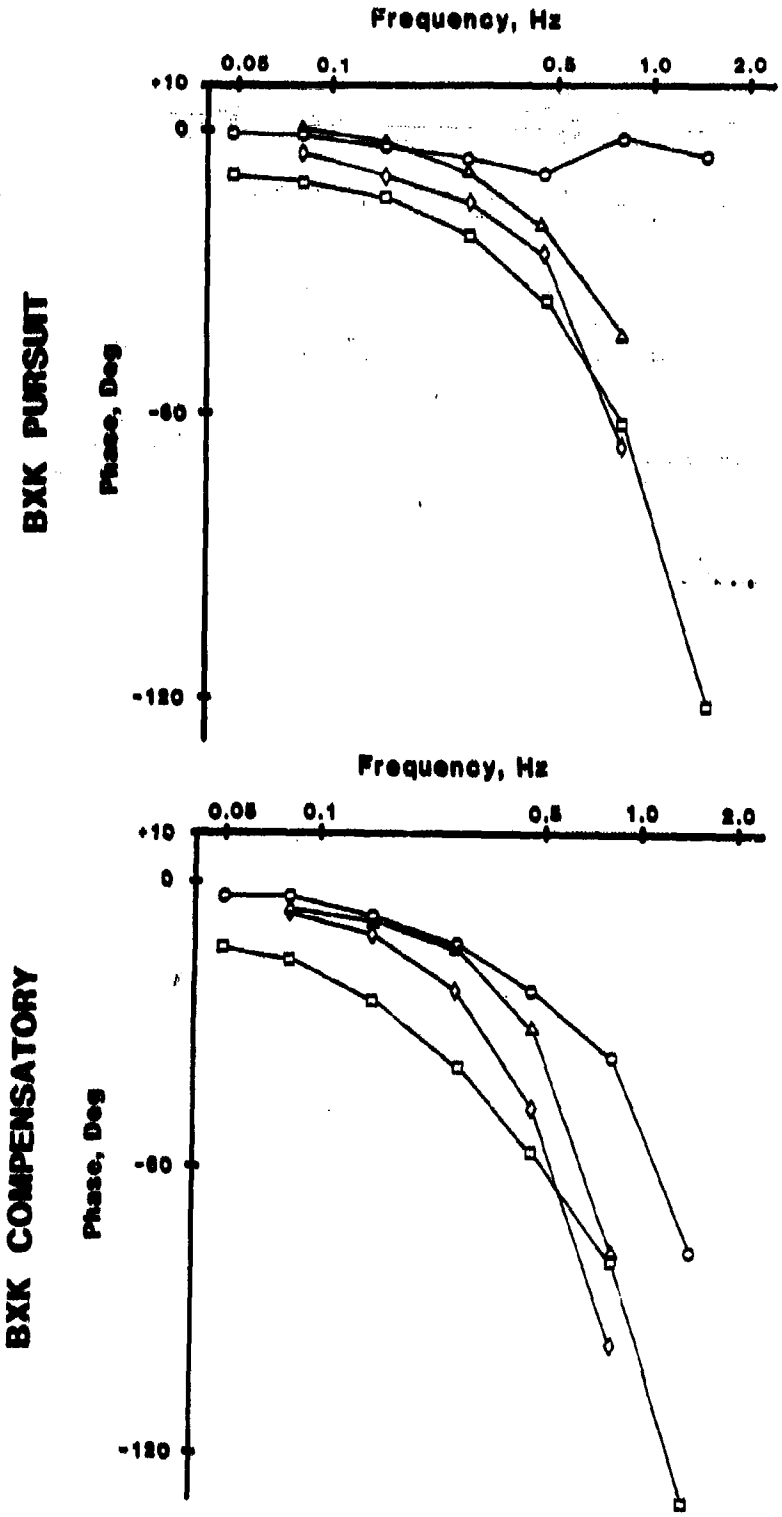

Frequency, Hz
0.05
0.1
0.5
1.0
2.0
+10
0
-60
-120
BXK PURSUIT
Phase, Deg
Frequency, Hz
0.05
0.1
0.5
1.0
2.0
+10
0
-60
-120
BXK COMPENSATORY
Phase, Deg

Figure 5

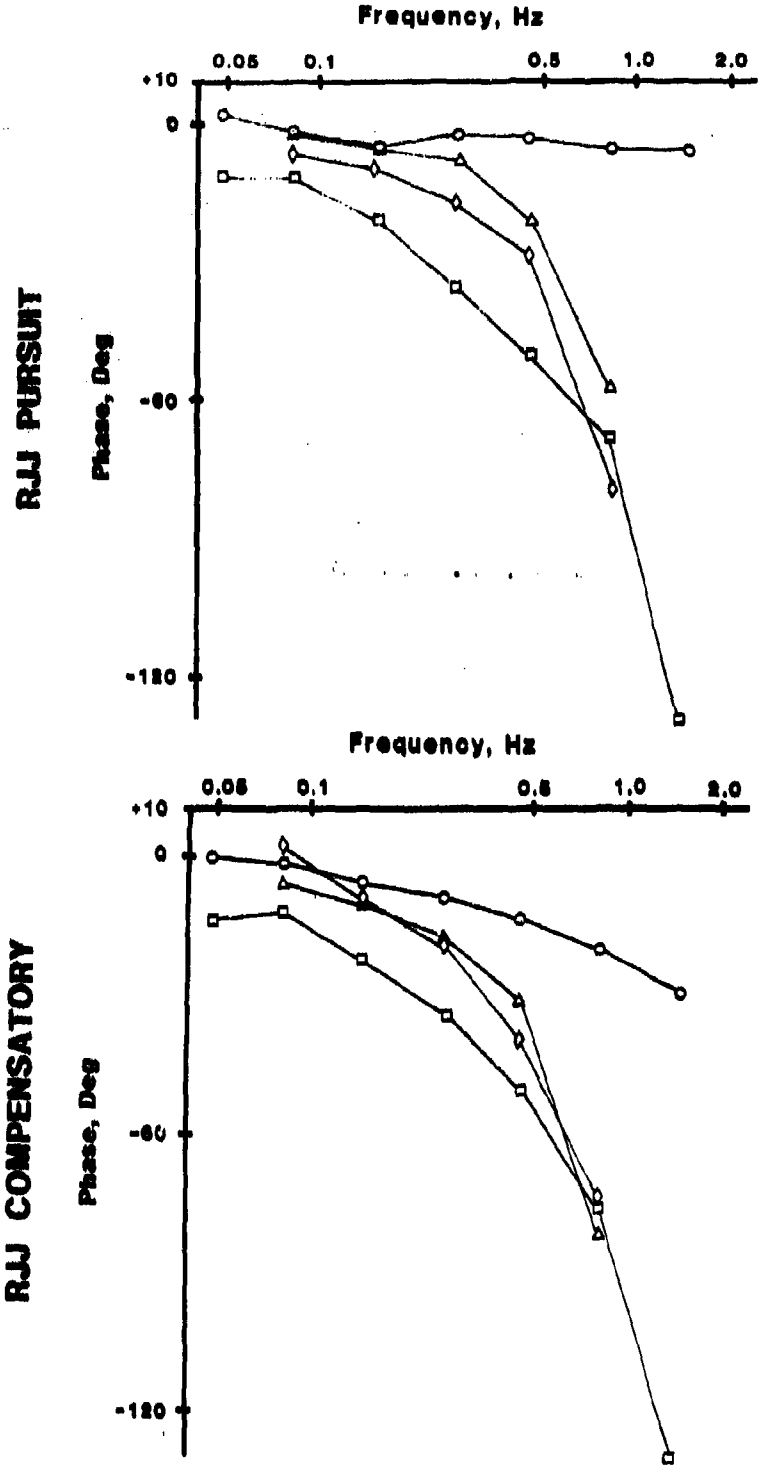

Frequency, Hz
0.05
0.1
0.5
1.0
2.0
+10
0
-60
-120
RJJ PURSUIT
Phase, Deg
Frequency, Hz
0.05
0.1
0.5
1.0
2.0
+10
0
-60
-120
RJJ COMPENSATORY
Phase, Deg

Figure 6

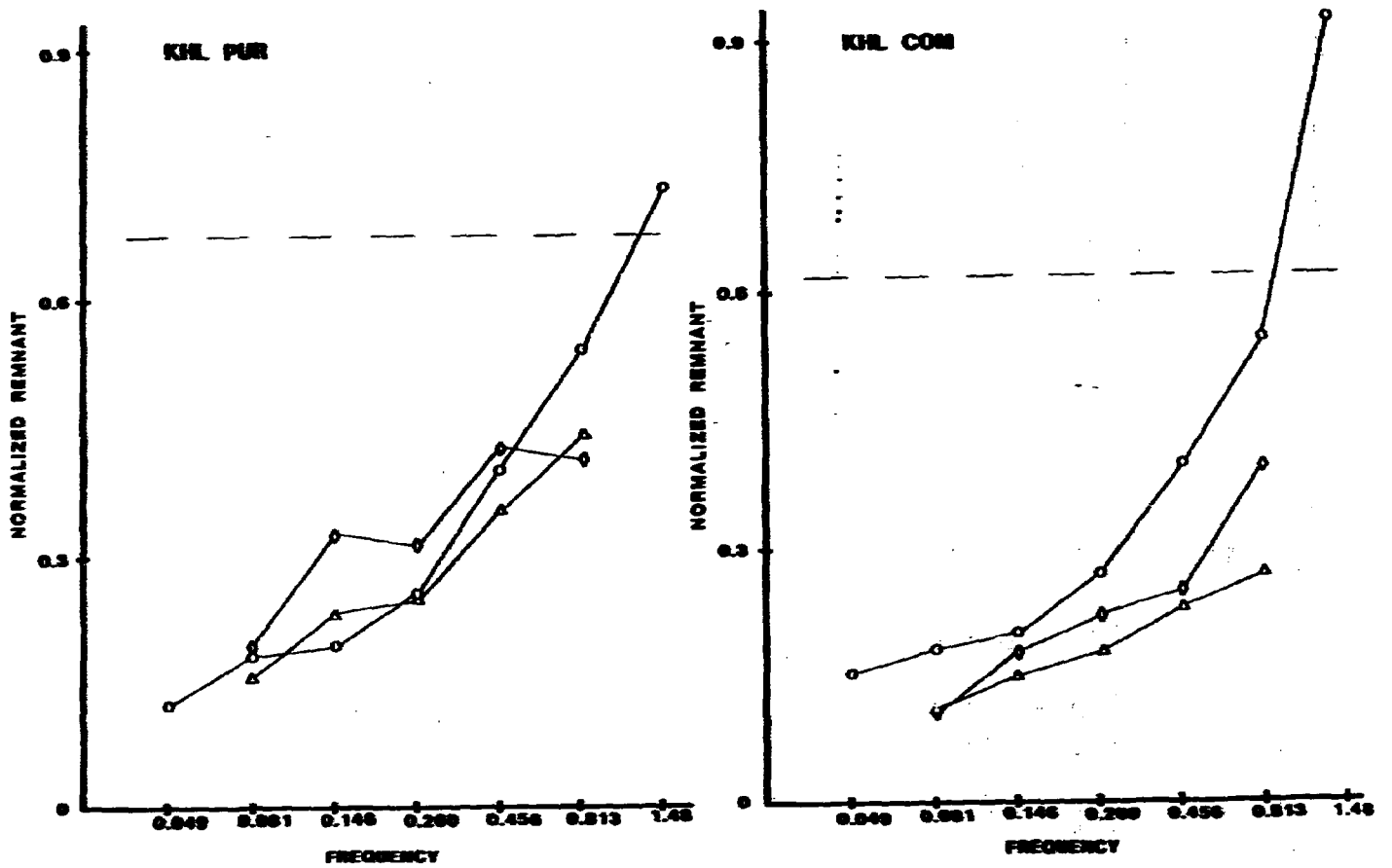

KHL PUR
KHL COM
NORMALIZED REMNANT
FREQUENCY
0.9
0.6
0.3
0
0.849
0.981
0.146
0.299
0.456
0.813
1.48

Figure 7

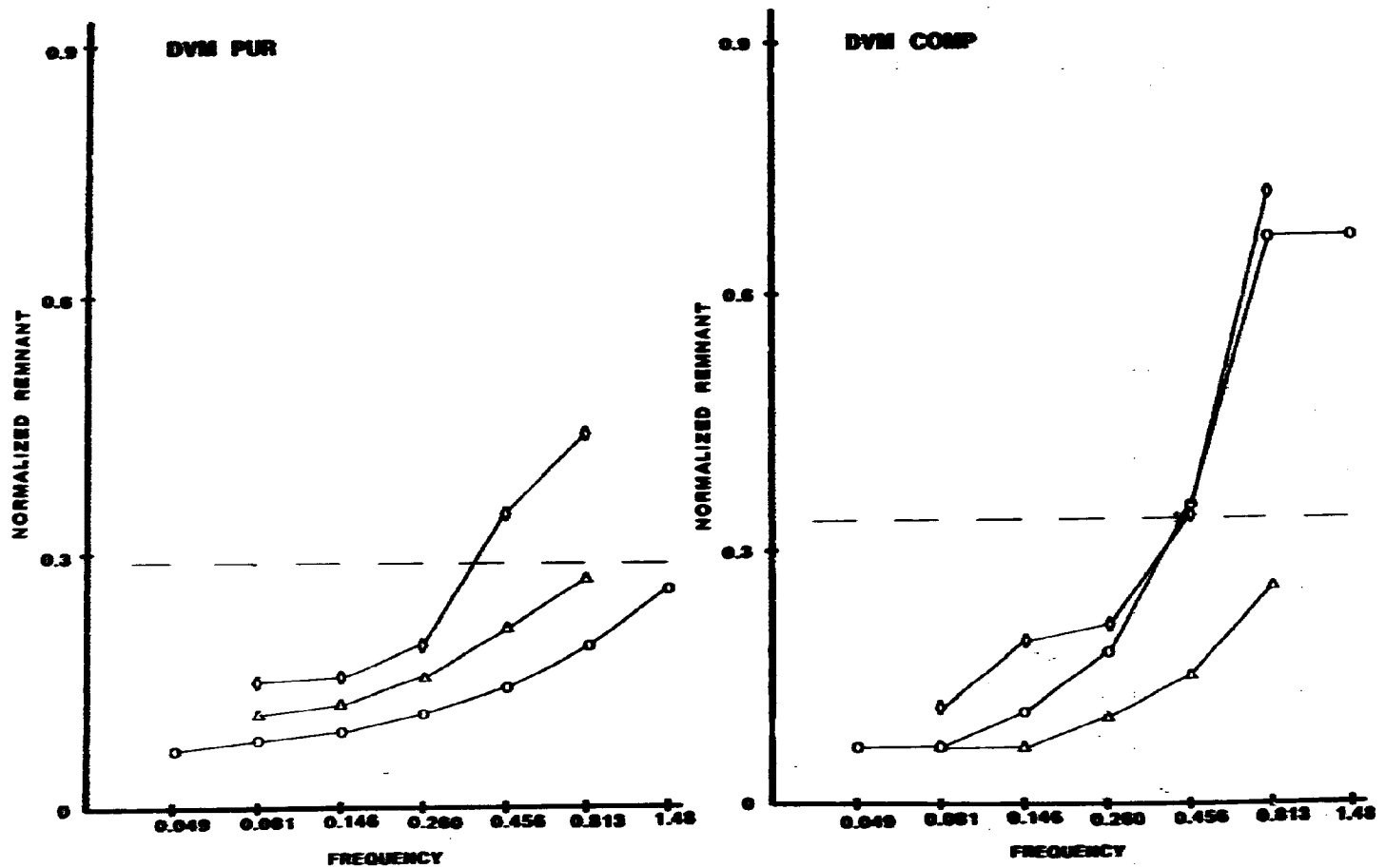

DVM PUR
NORMALIZED REMNANT
0.9
0.6
0.3
0
0.049
0.081
0.146
0.266
0.456
0.813
1.48
FREQUENCY
DVM COMP
NORMALIZED REMNANT
0.9
0.6
0.3
0
0.049
0.081
0.146
0.266
0.456
0.813
1.48
FREQUENCY

Figure 8

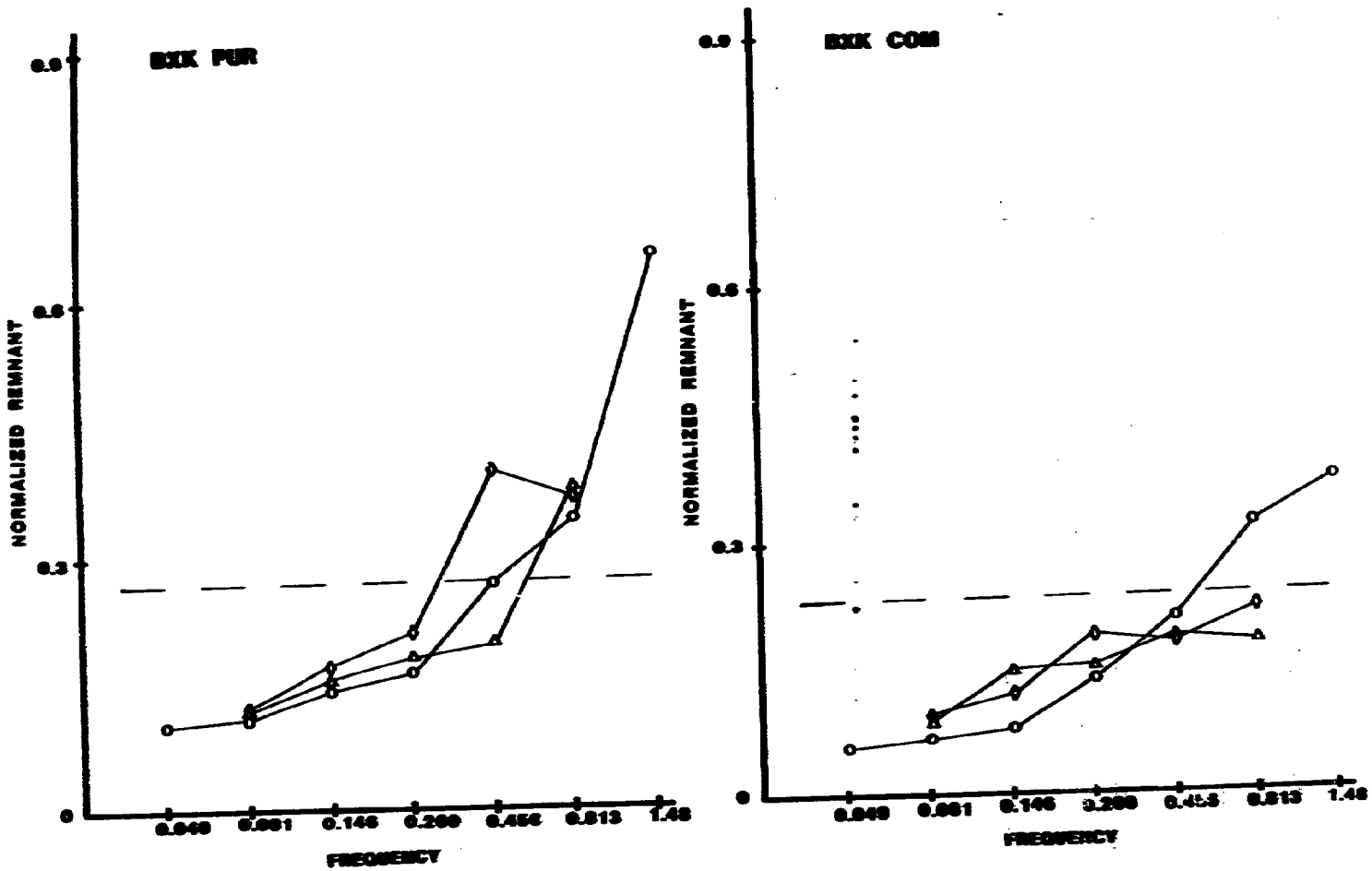

BXK PUR
NORMALISED REMNANT
0.9
0.6
0.3
0
0.146
0.200
0.813
1.48
FREQUENCY
BXK COM
NORMALISED REMNANT
0.9
0.6
0.3
0
1.48
FREQUENCY

Figure 9

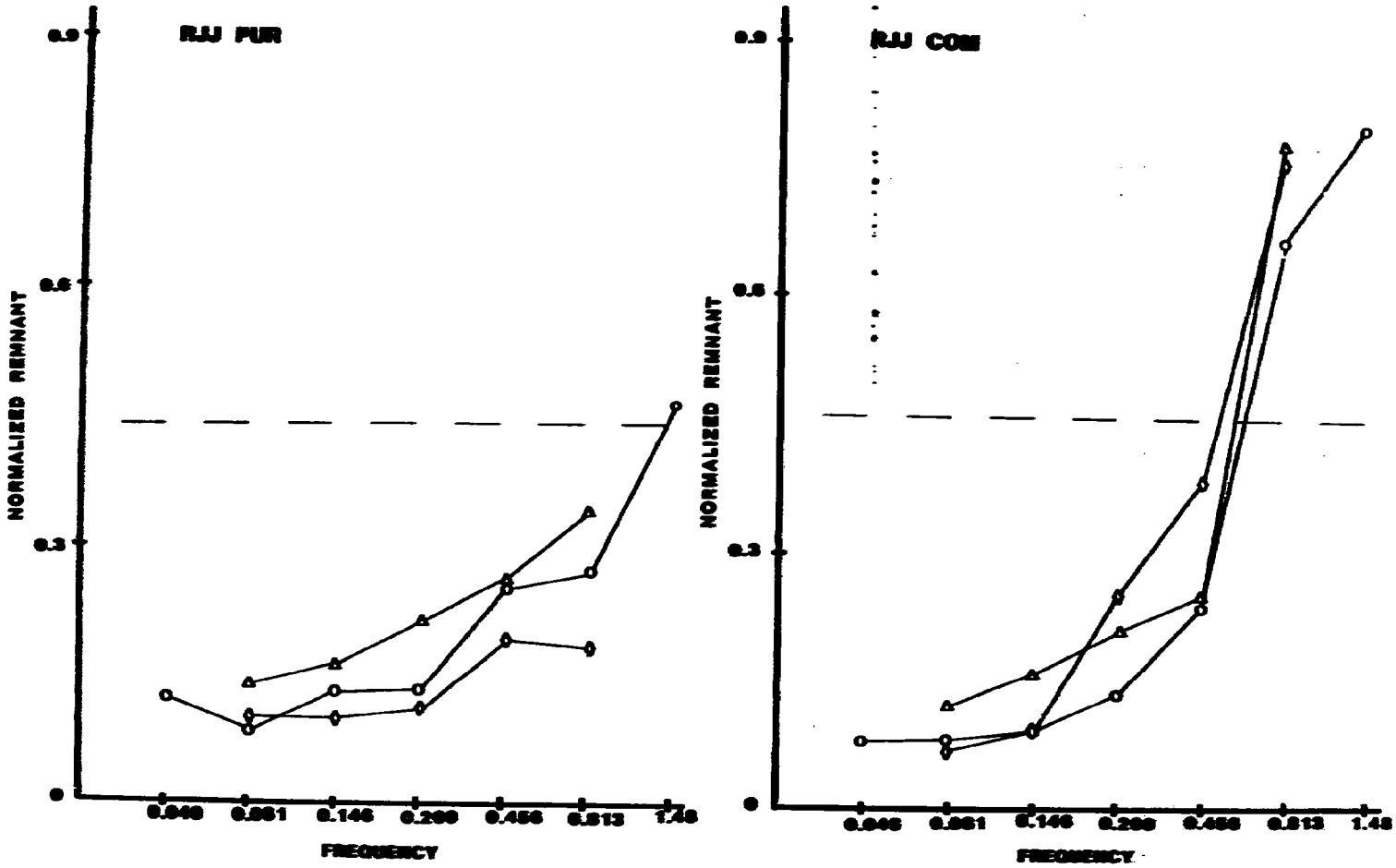

R.U PUR
R.U COM
NORMALIZED REMNANT
0.9
0.6
0.3
0
0.046
0.081
0.146
0.266
0.456
0.813
1.46
FREQUENCY

Figure 10

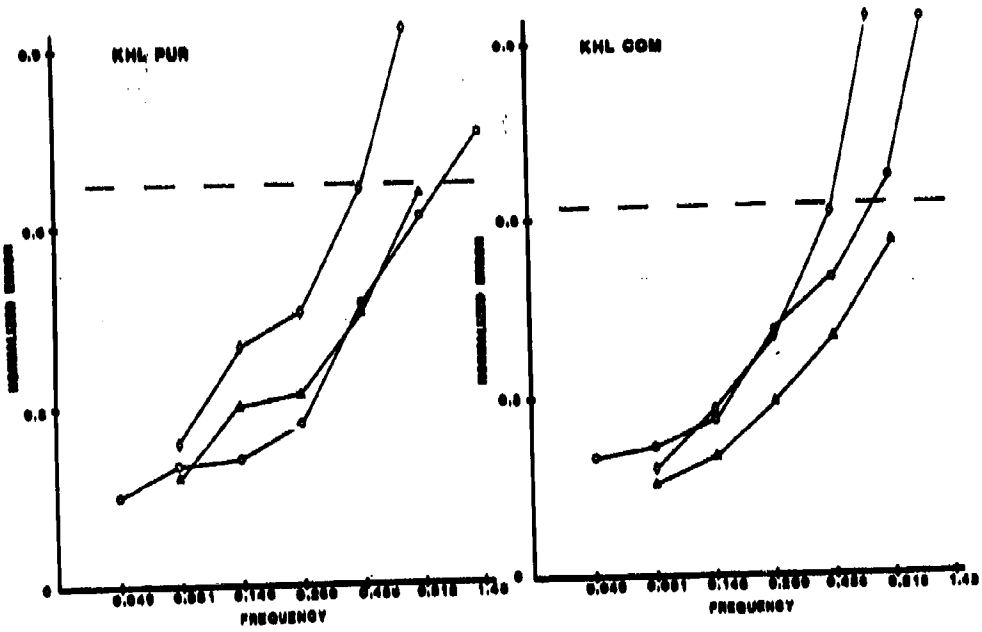


Figure 11

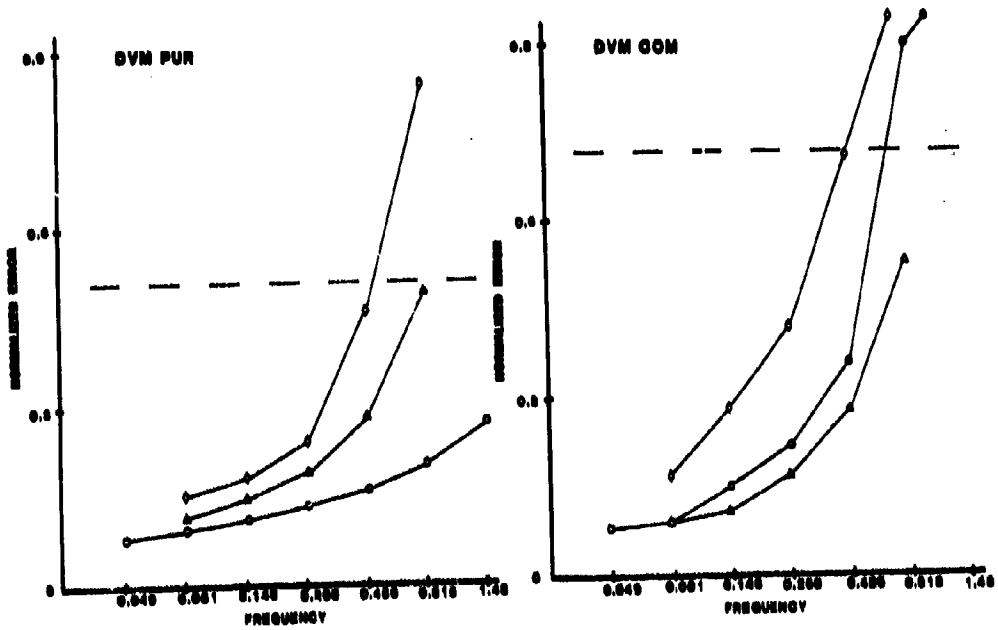


Figure 12

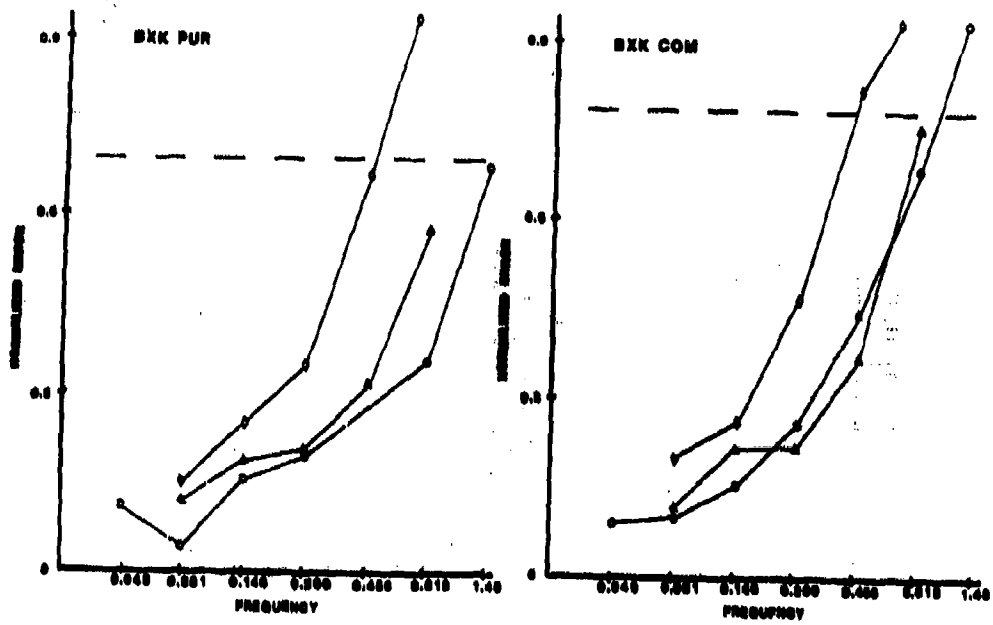


Figure 13

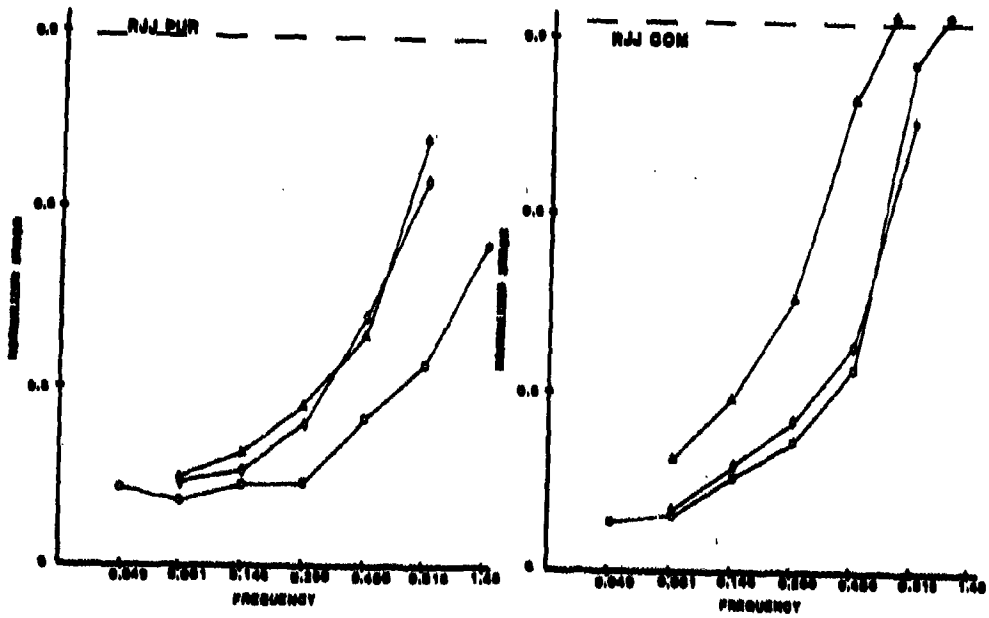


Figure 14

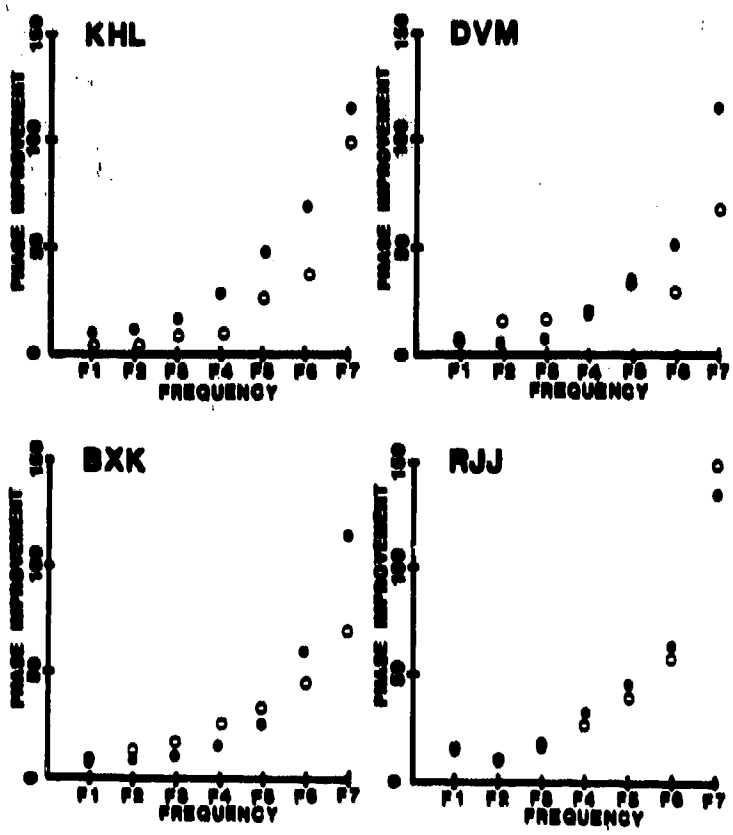


Figure 15

ADVANCES IN MODELING PILOT TRACKING PERFORMANCE IN THE PRESENCE OF SUSTAINED LINEAR ACCELERATIONS

by

Jonathan Korn and David L. Kleinman
Department of Electrical Engineering
and Computer Science
University of Connecticut
Storrs, Connecticut 06268

Abstract

Simulated air-to-air compensatory tracking experiments have been conducted at AMRL. These experiments involved closed-loop centrifugation of human subjects, i.e. the pilots experienced G-stress whenever pitch rate commands were applied. The multi-subject ensemble statistics clearly show that high G levels and transitive G periods have degrading effects upon pilot performance. These effects are quite well reproduced by the Optimal Control Model of human response, modified to include high-G phenomena.

I. Introduction

A distinctive characteristic of an aerial tracking task is the sustained positive acceleration, denoted as $+G_z$, that acts on the pilot. This is of paramount concern in high-performance aircraft during rapid turns and pullouts from dives. In the past, many experiments were devised to evaluate the degradation in pilot performance in a high G stress environment. In later works, the problem of analytically evaluating pilot performance has been considered. Therefore, a model of human performance under G-stress has become a developmental goal.

Korn and Kleinman [4] obtained a preliminary model of pilot performance that employed the Optimal Control Model (OCM) [2]. In a further work, Korn *et al.* [3] approached the modeling effort by assuming time varying control strategies and G-dependent OCM parameters. The data base for the development and validation of these models has been generated on the Dynamic Environment Simulator centrifuge (DES) at the Aerospace Medical Research Laboratories (AMRL), WPAFB. This paper is an extension of the work reported in Reference [3]. The following sections describe the experimental and modeling phases.

II. The Experiment

Repetitive trials of a compensatory tracking task were conducted on the DES facility under two experimental conditions. In both situations, the subjects were seated in the centrifuge cab and were required to track a triangular-shaped aircraft image presented to them on a graphics screen. In the first experimental condition, the centrifuge was at rest, whereas in the second it was engaged in a closed-loop operation. Thus we use the terminology static- vs. dynamic-G. The target motion followed a predetermined pattern in the longitud-

inal (pitch) axis only. The subjects, responding to an error stimulus, generated a commanded control input in order to reduce the tracking error. By this action, they induced a positive, time varying G_z stress upon themselves by increasing or by decreasing the angular velocity of the centrifuge (in the dynamic G runs).

Six subjects participated in these experiments. The particular experiments that were of concern to this study consisted of thirty-eight replications of a 60-second static-G tracking run, and thirty-six replications of the dynamic-G run.

From the recorded longitudinal and lateral tracking error ensemble, and the subject attained G profile (G_A), first- and second-order statistics were computed. In the present study we will be concerned with the lateral axis data[†] that was found to exhibit the following characteristics:

1. The lateral tracking error ensemble mean , $\bar{a}(t)$, obtained from the dynamic-G runs is not statistically different from the static-G mean. In both cases it is virtually zero as might be expected---there was no target input in the horizontal axis. The tracking error in this axis is induced by the pilot's erratic control inputs, e.g. inherent motor randomness, axis crossfeed, etc.

2. The ensemble standard deviation, $\sigma_a(t)$, for the dynamic-G condition is significantly larger ($P<.01$) than that of the static-G case (Figures 3 - 4). The F-test showed a significant difference ($P<.01$) between the static and dynamic-G variance time histories over nearly the entire 60-second period. The differences are most pronounced during the peak G_A and during high $\dot{G}_A$ periods, as can be seen by comparing Figures 3 and 4 with Figure 5. This clearly indicates that there is a substantial degradation in lateral tracking performance when operating under G-stress.

A descriptive model for this performance degradation is developed in the subsequent sections.

III. The System

The lateral tracking task represents essentially a side-task as there are no target aircraft motions in the horizontal plane. In order to model the lateral task, the dynamics between aileron deflection, δ_a, and tracking error, a, must be defined. The modeling process assumes the following:

1. Roll angle, ϕ, is small so that $\cos\phi \simeq 1$ and $\sin\phi \simeq \phi$. This enables us to treat the lateral and longitudinal modes as uncoupled. Moreover, the longitudinal tracking task is not dependent on ϕ in this case.

2. The attacker angle-of-attack α_A and sideslip angle $\beta_A = 0$. Thus, the attacker velocity vector is always aligned with the body axis. This greatly simplifies the system representation in the horizontal plane.

3. All turns are coordinated, i.e. the G vector is aligned with the aircraft's z-body axis.

[†] Logitudinal data analysis and the related modeling work are described in Ref. [3].

For a level, coordinated, turn

$$\dot{\Psi} = g \frac{\phi}{V} \tag{3.1}$$

and when the aircraft is pitching

$$\dot{\Psi} = (q + \frac{g}{V}) \phi \quad ; g = 32.2 \text{ ft/sec}^2 \tag{3.2}$$

where Ψ, q, and V are aircraft heading angle, pitch rate and velocity, respectively. By the above assumptions $q \approx \dot{\gamma}$ for the vertical axis, γ being the flight-path angle. Thus

$$\dot{\Psi}_T \approx (\dot{\gamma}_T + \frac{g}{V}) \phi_T$$

and (3.3)

$$\dot{\Psi}_A \approx (\dot{\gamma}_A + \frac{g}{V}) \phi_A$$

where the subscripts T and A denote the target and the attacker, respectively. Since we are interested only in the deviations between attacker and target aircraft motions, it is convenient to consider an attacker centered coordinate system. Thus, we set $\phi_T = 0$ and $\Psi_T = 0$. It is convenient to define

$$\begin{aligned} \phi &\triangleq \phi_A - \phi_T = \text{relative roll angle} \\ \Psi &\triangleq \Psi_A - \Psi_T = \text{relative headings} \end{aligned} \tag{3.4}$$

so that,

$$\dot{\Psi} = (\dot{\gamma}_A + \frac{g}{V}) \phi \tag{3.5}$$

The time-varying quantity $\dot{\gamma}_A$ affects the motion in the lateral axis. In fact, if we consider $\dot{\gamma}_A$ = constant then Equation (3.5) represents a dynamic linearization. In the present problem $\dot{\gamma}_A$ is a random variable, with a mean and variance that are functions of time. It is more convenient to use $\dot{\gamma}_T$ in Equation (3.5), since this will be a "cleaner" signal. The approximation is valid to first-order as $(\dot{\gamma}_T - \dot{\gamma}_A) \phi \approx 0$. Thus,

$$\dot{\Psi} = (\dot{\gamma}_T + \frac{g}{V}) \phi \tag{3.6}$$

The next equation specifies the relationship between heading error Ψ and tracking error e(t). It is easy to show that

$$\dot{e} = \frac{V}{D} \Psi \tag{3.7}$$

where D = distance between the two aircraft.

The final equation reflects the aircraft roll axis dynamics which is essentially a roll-rate command system

$$\tau_\phi \ddot{\phi} + \dot{\phi} = \delta_a \tag{3.8}$$

In the present simulations $\frac{1}{\tau_\phi} = 5.55$.

Equations (3.6) - (3.8) show that the response between δ_a and e(t) is like $1/s^3$. This is a very difficult system for a human to control, depending on the visual information. The block diagram of this system is shown in Figure 1.

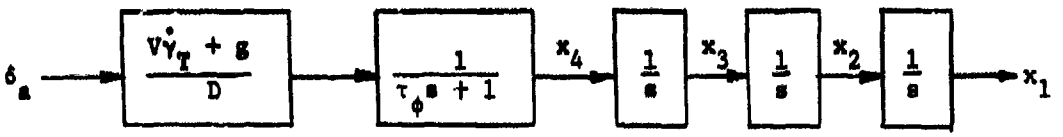


FIGURE 1. A 4th ORDER LATERAL AXIS MODEL

For convenience, the scale factor, or open-loop gain, $(V\dot{\gamma}_T + g)\ D^{-1}$ has been made into an equivalent gain K(t) at the input†. As a result, the state definitions become

$$\begin{aligned} x_4 &= K(t)\ \dot{\phi} \\ x_3 &= K(t)\ \phi \\ x_2 &= \dot{a}(t) = \frac{V}{D}\ \psi \\ x_1 &= a(t) \end{aligned} \tag{3.9}$$

It should be noted that the above manipulation is valid only when <u>K(t) is constant or slowly varying</u>. This has been the usual assumption when considering time-varying system parameters in pilot modeling efforts.

The system of Figure 1 is 4^{th}-order. It can be reduced to third order by combining the first-order lag $(\tau_\phi s + 1)^{-1}$ with the first order lag $(\tau_N s + 1)^{-1}$ that is introduced by the neuro-motor dynamics in the optimal control model. Thus we set $\tau_{N,equiv} = \tau_N + \tau_\phi$ in the subsequent modeling process, and the system of Figure 1 becomes as shown in Figure 2.

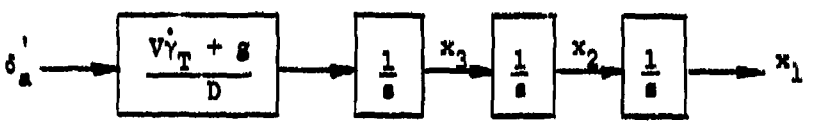


FIGURE 2. A 3rd ORDER LATERAL AXIS MODEL

Here, $\delta_a' = \dot{\phi}$ and x_1, x_2, x_3, are as defined previously. The state equations are

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= K(t)\ \delta_a' \end{aligned}$$

† Note $V\dot{\gamma}_T + g$ is the target acceleration.

or, in matrix form,

$$\dot{\underline{x}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \underline{x} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} K(t)\ \delta_a'(t) = A\underline{x} + \underline{b}K(t)\ \delta_a'(t) \qquad (3.10)$$

Note that K(t) enters only as a scalar multiplier on the control δ_a'. This means that the control δ_a' can be computed for K(t) = 1 and then adjusted for arbitrary K(t) via inverse scaling. It will be seen that K(t) has an effect on the pilot's additive motor noise. For notational purposes we define $K(t)\delta_a'(t) = u(t)$.

The observations that are available to the pilot in order to minimize lateral tracking error are

$$\begin{aligned} y_1 &= x_1 && = \text{tracking error } e(t) \\ y_2 &= x_2 && = \text{error rate, } \dot{e}(t) \\ y_3 &= x_3/K(t) && = \text{relative roll angle, } \phi \\ y_4 &= u/K(t) && = \text{roll rate, } \dot{\phi} \end{aligned} \qquad (3.11)$$

The observation y_3 of roll angle is critical for the control task. If this information is absent it is virtually impossible for a human to control the K/s^3 system. The rule-of-thumb control logic of "keep your wings aligned with those of the target aircraft" is testimony to this fact. The information y_4 is not very important here inasmuch as the control signal u(t) is essentially "known". It would be of importance if the target aircraft were free to move laterally.

The above developments specify completely the equations needed to apply the optimal control model as presented in the next section.

IV. The Optimal Control Model [2].

The operator-adopted tracking strategy is represented via the OCM cost functional weighting parameters. These weightings are used to generate "optimal" feedback controls. But the human operator has inherent limitations that limit his performance in a control task. These limitations are represented in the OCM by perceptual and/or motor submodels that include several parameters. It is our hypothesis that the human's limitations and tracking strategy are affected by G-stress. Therefore, the modeling approach is to first isolate the G-stress dependent OCM parameters and cost functional components, and then to determine the structure of that dependency. Below we list all OCM parameters and indicate how they are affected by the acceleration stress.

Observation / Indifference Thresholds: The pilot's perceived observations $y_{pi}(t)$, are a noisy, delayed and thresholded version of the displayed variables, viz,

$$y_{pi}(t) = F_i[y_i(t-\tau)] + v_{yi}(t-\tau) \qquad i=1,\ldots,4 \qquad (4.1)$$

Here $y_i(\cdot)$ is the displayed variable, $v_{yi}(\cdot)$ is a white observation noise, τ is the nominal lumped time delay (.2 sec) and $F_i(\cdot)$ is the threshold nonlinearity of width $\pm a_i$. In the context of this study, the a_i's (i=2, 3, 4) were selected at nominal values. The indifference threshold on $y_1 = e$, however, was treated differently.

The experimental setup was such that the subjects were instructed to minimize the tracking error only during the second half of the tracking period. Therefore, during the initial period (including the high G peak- see Figure 5) the subjects adopted a strategy of merely "keeping the target on the screen". This variation in pilot control strategy could be modeled by altering the indifference threshold on the tracking error. Thus, during the first tracking inverval ($0 \le t \le T_1$, T_1 = 30 sec) a_1 = 80 mrad = maximum angular error allowable, given the display screen width. During the second half of the run we choose a_1 = 30 mrad corresponding to a minimal threshold level (= 1/3 pipper diameter).

The increase in error threshold described above is independent of G forces. In the presence of G-stress, there is assumed to be a further increase in the model's indifference threshold. The argument is that under sustained physical stress the pilot tends to decrease his concern over the tracking error. In other words, he becomes increasingly indifferent to the information presented on the display screen, and more concerned with his own psycho-physiological response. Assuming a minimal G_A level (ΔG_{min} = 4g) below which the pilot would perform "well", we modify the indifference threshold according to the heuristic formula,

$$a_1(t) \rightarrow a_1(t) \cdot \left(1 + c \cdot \max^2 [0, (G_A - G_{min})]\right) \tag{4.2}$$

where c is an adjustable constant. In the present work c = .2.

In the OCM it is assumed that the observations $y_i(t)$ are normally distributed random variables. The Random-Input-Describing-Function (RIDF) is conveniently used to statistically linearize $F(\cdot)$. The perceived observations, Eq (4.1), are therefore approximated by

$$y_{pi}(t) = N_i(a_i, \bar{y}_i, \sigma_{yi})\, y_i(t-\tau) + v_{yi}(t-\tau) \qquad i=1,\ldots,4 \tag{4.3}$$

where $N_i(\cdot)$ is the equivalent linear gain and $\bar{y}_i$ and σ^2_{yi} are the mean and the variance of the i-th observed variable.

Motor Noise: In the Optimal Control modeling process, the control δ_a' is generated as

$$\tau_{N,equiv}\, \dot{\delta}_a' + \delta_a' = -L(t)\underline{x}(t) + v_m(t) \ . \tag{4.4}$$

The gains L(t) are computed via

$$L(t) = L_1/K(t) \tag{4.5}$$

where L_1 are the optimal gains corresponding to K(t) = 1. The quantity $v_m(t)$ is the human's "motor" noise. The motor noise consists of two parts: a purely additive component $v_m^o(t)$ and a multiplicative component that scales with the control δ_a'. Thus,

$$v_m(t) = v_m^o(t) + \rho_u |\delta_a'(t)| \xi(t) \tag{4.6}$$

where $\xi(t)$ is assumed to be "white" and ρ_u is the motor noise ratio. Multiplying Equation (4.4) through by K(t), noting Equation (4.5) and defining $\delta_a' K(t) \triangleq u(t)$, we obtain

$$\dot{\underline{x}} = A\,\underline{x} + \underline{b}\, u(t) \tag{4.7}$$

$$\tau_{N,equiv}\dot{u} + u = -L_1\underline{x}(t) + K(t)v_m^o(t) + \rho_u |u(t)| \xi(t) \tag{4.8}$$

Thus, the only effect of the gain K(t) is to increase the effective <u>additive</u> motor noise component, the variance of which increases as $K^2(t)$, viz

$$V_m^o(t) + K^2(t)\ V_m^o(t)$$

The <u>multiplicative</u> part of the motor-noise scales with the effective control u(t), which is convenient for modeling. The net result is as expected: A constant motor noise input will have a greater effect on lateral tracking error during periods of high pitch rate since the system sensitivity, K(t), increases as $q = \dot{\gamma}$ increases.

The above discussion is relative to the interdependencies between the system dynamics and the motor noise. The effect of G-stress on this motor noise remains to be determined. In the course of the modeling work, it became apparent that model predictions were quite sensitive to variations in V_m^o and ρ_u. Increases in these parameters greatly increased the magnitude of the model-predicted $\sigma_e(t)$. Large increases in σ_e are observed between the static and the dynamic-G conditions (Figures 3-4). We postulate therefore that G and/or $\dot{G}$ stress increases the motor-noise. In a previous effort to model pilot performance under vibration stress, Levison [5] also needed to increase the motor noise to match the data. Although vibration is a different form of physical stress than sustained acceleration, it is possible to extrapolate from Levison's results to our study. The following approach was taken.

1. In the static conditions the nominal value of $V_m^o = .15$ was chosen on the basis of the manipulator dead-zone characteristics. In the dynamic conditions $V_m^o = 1$.

2. Attained-G levels increase $V_m^o(t)$ according to a relation of the form,

$$V_m^o(t) = V_m^o \left(1 + c \cdot \max^2 [0, (G_A - G_{min})]\right)$$

where $G_{min} = 4g$ and $c = 0.2$ is chosen to match the σ_e data.

3. Sharp peaks occur in the ensemble error standard deviation in Figure 4 at t = 20 sec. and t = 35 sec. At these times pilot performance is the worst. Also, from Figure 5, it is clear that $|\dot{G}_A|$ is at its highest levels around these times. It would appear that these two events are correlated, and that the data trends can be explained by the crossfeed phenomenon in a manipulator, enhanced by the $\dot{G}_A$ factor. We therefore apply $V_m^o = 400$ (i.e. 20 times higher additive motor noise in the lateral axis) during the intervals $12 \leq t \leq 21$ and $30 \leq t \leq 36$, where $|\dot{G}_A|$ is maximum. This is, of course, only a tentative value and an appropriate functional relationship $V_m^o(t) = V_m^o(t, G_A, \dot{G}_A)$ has to be determined.

4. The nominal value of ρ_u is -20dB for the static modeling condition and -18dB for the dynamic case.

Observation Noise: This noise was chosen in the usual way

$$V_{yi} = \rho_{yi} \frac{\sigma_{yi}^2}{f_i} \qquad i=1,\ldots,4 \qquad (4.9)$$

where V_{yi} is the observation noise variance, ρ_{yi} is the observation noise ratio (nominal value of -20dB) and f_i is the fractional attention allocation to indicator i (=1/4 for all i)

Neuro-Motor Time Constant: The effective time constant used in this study was

$$\tau_{N,equiv} = \tau_N + \tau_\phi = .1 + \frac{1}{3.33} = .3$$

Performance Index: The optimal control policy is that which minimizes the cost functional

$$J(u) = E\left\{q_e e_p^2 + q_\phi \phi_p^2 + g\dot{u}^2\right\} \qquad (4.10)$$

The q_e, q_ϕ and g are weighting constants, and e_p and ϕ_p are the <u>perceived</u> observations of the lateral tracking error and relative roll angle †. In effect, the performance index becomes

$$J(u) = E\left\{N_1^2[a_1(t),\ \sigma_e(t)]\cdot q_e e^2 + q_\phi \phi^2 + g\dot{u}^2\right\} \qquad (4.11)$$

(The relative roll angle indifference threshold is negligible (a_3=0), so that $\phi=\phi_p$.) With this modification, the cost functional weighting e(t) is no longer a constant, but rather a time-varying quantity since the threshold a_1 and the error standard deviation σ_e are time functions. This, in turn, results in time-varying feedback control gains.

V. Results and Conclusions

Once the OCM parameters are determined, as in the previous section, the model is completely specified. Application of the model does indeed reproduce the experimental data, as shown in Figures 3 and 4. The data-model fits are excellent considering the fact that no effort has yet been directed towards a fine-tuning identification of the pertaining OCM parameters. A discussion of the results follows.

<u>Static-G</u> (Figure 3): The data exhibit a slight "transient" that apparently results from the effects of coupling between vertical and horizontal axes. There are no other distinctive trends in the data throughout the entire 60 second period. The level of the ensemble standard deviation decreases from about 30 mrads at the "transient" peak to about 20 mrads in the steady state. It is evident that the model reproduces the experimental data very well. It should be emphasized that no modifications to the OCM nominal parameters were necessary in the static condition. This reaffirms the validity of the OCM when nominal conditions are considered.

<u>Dynamic-G</u> (Figure 4): Examination of the experimental data reveals the following.

1. There is a large "transient" in the ensemble error standard deviation of up to 70 mrads.

2. The standard deviation eventually levels off to about 25 mrads.

3. Two sharp peaks are observed around t=20 and t=35 seconds, apparently due to a dramatic increase in the pilot motor randomness during transitive G_A periods.

The reproduction of data by the model is very good. The "transient" level and the steady-state recovery are well-replicated and the two distinct peaks are

† The use of the <u>perceived</u> rather than the <u>displayed</u> variables in the cost functional was first suggested in Ref. [1].

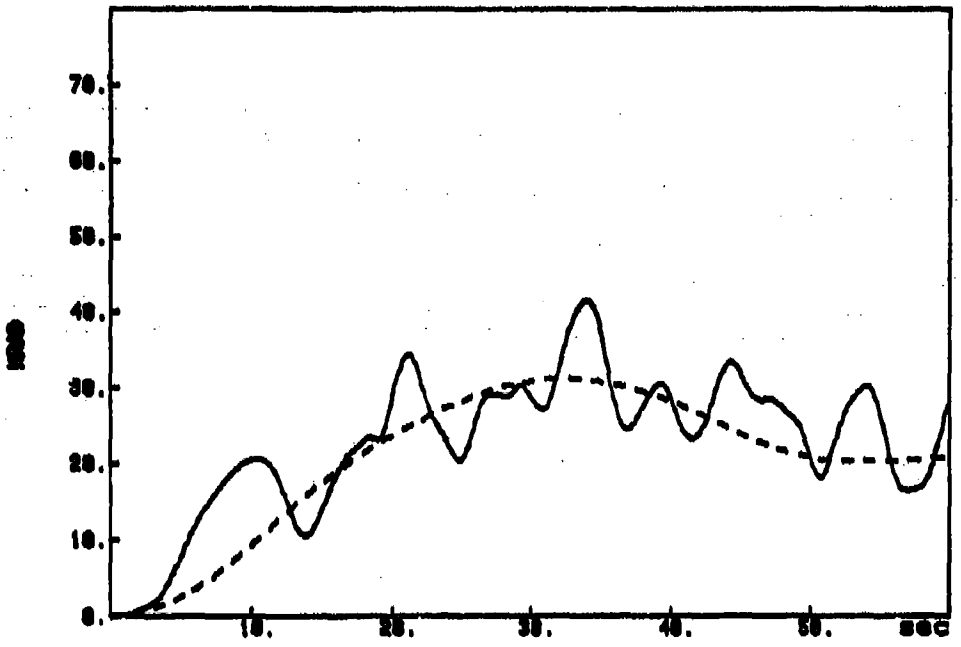


FIGURE 3
LATERAL TRACKING ERROR - ENSEMBLE STANDARD DEVIATION
STATIC G

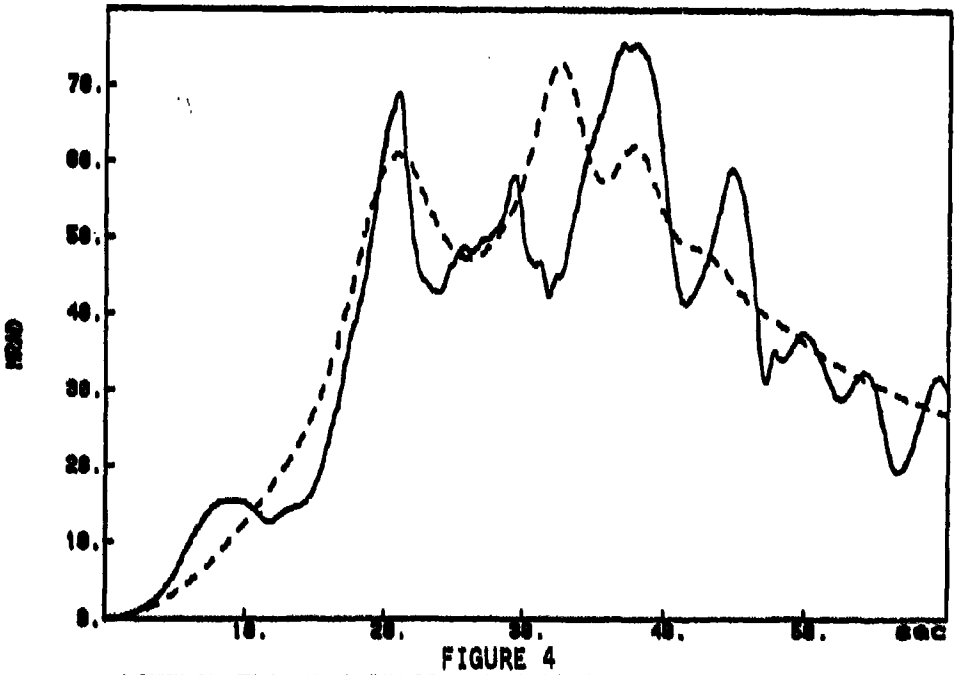


FIGURE 4
LATERAL TRACKING ERROR - ENSEMBLE STANDARD DEVIATION
DYNAMIC G

predicted. The model also shows an additional peak around the t=32 second which seems to be somewhat misplaced if not outright puzzling.

In conclusion, it must be re-emphasized that the model presented in this paper is by no means in final form. One may question the significance of the various trends in the data. For example, it is questionable whether the sharp peak in $\sigma_e(t)$ at t=20 for the dynamic-G condition (Figure 4) is significant. If we deem that this trend is not significant, no special increase in the additive motor noise in this region should be assumed. In such case, the model would predict no peak in this interval. It is generally difficult to avoid these types of problems when identifying models in man-machine systems, especially if time-domain data are analysed. Thus additional experiments and further model refinement are desired. These efforts are now underway.

References

1. Kleinman, D. L. and Baron, S., Bolt Beranek and Newman, Inc., Cambridge Mass., "Analytic Evaluation of Display Requirements for Approach to Landing," March 1971, BBN Report 2075.

2. Kleinman, D. L., Baron, S. and Levison, W. H., "A Control Theoretic Approach to Manned-Vehicle Systems Analysis," IEEE Transactions on Automatic Control, Volume AC-16, No. 6, 1971.

3. Korn, J., Boal, H. S. and Vikmanis, M. M., "Modeling Human Tracking Performance in a High-G Stress Environment," Proceedings of the 17th. Conference on Decision and Control, January 1979.

4. Korn, J. and Kleinman, D. L., "Modeling the Effects of High-G Stress on Pilots," Proceedings of the 14th. Annual Conference on Manual Control, 1978.

5. Levison, W. H., "Analysis of Vibration-Induced Pilot Remnant," Proceedings of the 10th. Annual Conference on Manual Control, 1974.

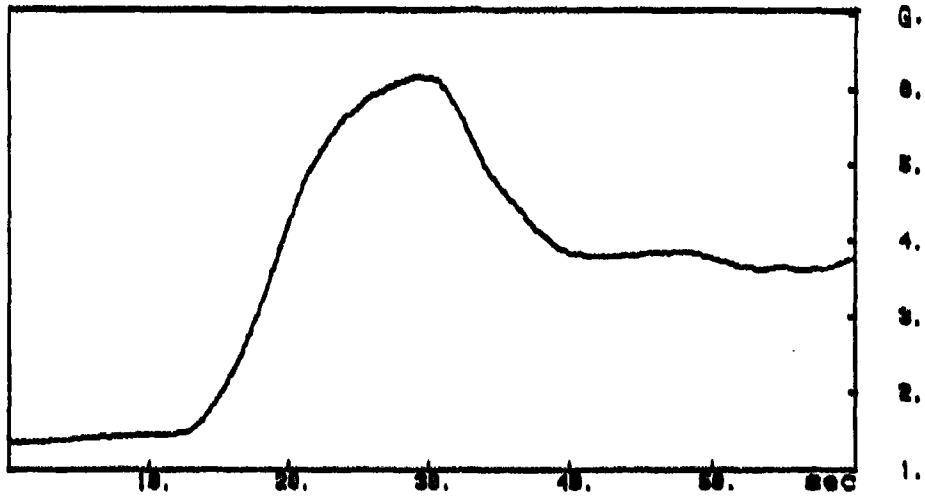


FIGURE 5 PILOT ATTAINED G - ENSEMBLE MEAN

INVESTIGATION OF ALTERNATE HUMAN OPERATOR OPTIMAL CONTROL MODEL STRUCTURES*

- Anil V. Phatak -
Senior Engineer
Analytical Mechanics Associates, Inc.
Mountain View, CA 94043

ABSTRACT

The standard optimal control model (SOCM) for the human operator as developed by Kleinman, Baron and Levison, incorporates a fixed set of hypotheses or assumptions in its formulation. These assumptions have been quite successful in describing human response behavior in a variety of control situations. However, within the stochastic optimal control theory (LQG) framework, several alternate sets of hypotheses appear capable of duplicating measured human response characteristics. This paper shows that there exist several alternate optimal control model structures in the class of optimal control models that are equally acceptable in describing standard laboratory data for k, k/s and k/s^2 plant dynamics. Implications of these findings on model structure selection and parameterization are discussed. Situations requiring additional experimental verification are identified. The paper also includes an evaluation of the identifiability properties of the optimal control model structures and their impact on the utility of such models for predictive purposes.

* This work was accomplished under contract No. F33615-78-C-0501 for the Air Force Aerospace Medical Research Laboratory through Request No. 22-ME-4 from the University of Dayton Research Institute, Dayton, Ohio.

I. INTRODUCTION

An important problem in manual control research is that of developing and validating a mathematical model (s) for the human operator involved in a realistic control task. The primary motivation behind this effort lies in the expected utility of a human operator model in the systematic analysis and prediction of manned system performance. A quantitative evaluation of the effects of the controlled system dynamics, the environmental disturbances, the display and control interface parameters and external or internal stressors on human operator control and performance is crucial if one is to achieve maximum performance and reliabilit, of the complete manual control system. Consequently, several different mathematical representations for the human controller have been proposed over the past thirty years. The human has been characterized as a servo compensator, a digital controller, a finite-state machine and an optimal controller. However, since its formulation a decade ago, the optimal control model of Kleinman, Baron and Levison [1-3], hereafter referred to as the Standard Optimal Control Model (SOCM), has emerged as one of the frequently used models for the study of complex man-in-the-loop systems. The primary reason for the success of the SOCM is the flexibility of the modern control theoretic method in handling multivariable, multiaxes, nonlinear and nonstationary stochastic control situations within a well developed and general state variable optimal control (Linear-Quadratic-Gaussian) framework.

The SOCM for the human operator incorporates a fixed set of hypotheses or assumptions in its formulation. These assumptions have been quite successful in describing human response behavior in a variety of control situations. However, within the stochastic optimal control theory framework, several alternate sets of hypotheses appear capable of duplicating measured human response characteristics. This report shows that there exist several alternate optimal control model structures in the general class of optimal control models that are equally acceptable in describing standard laboratory data for k, k/s and k/s^2 plant dynamics [5,6]. The implications of these findings on model structure selection and parameterization are substantial. The criterion for selection of one model structure over others must depend upon the intended use or application of the candidate model. If the intended use is to fit or mimic available experimental data, then clearly the model which gives the best fit or lowest residual variance must be chosen. A number of so-called "goodness of fit" tests have been developed to check whether the given model is a "good fit" to the data [7-9]. These tests are known as residual tests since the tests are given in terms of the residuals obtained from the model and the given set of observations. However, these significance tests have their limitations and can lead to erroneous conclusions if the intended model application includes prediction or forecasting in addition to data matching. Goodness of fit tests like any hypotheses test merely tell if the given model is completely inappropriate or not completely inappropriate (as opposed to completely appropriate). In other words the candidate model is either rejected or not rejected (as opposed to being accepted).

Thus if a model is to be used for prediction then matching available experimental data with minimum fit error is not a sufficient criterion for accepting (not rejecting) the model. Any human operator model selected should, as a minimum, be structurally identifiable and parsimonious in its parameterization. Unless the model structure is identifiable, there is no guarantee that the model parameters uniquely describe the measured human response data. In other words, different investigators can come up with several sets of model parameter values to describe the same input-output data. The need to have a model whose parameters are uniquely characterized by the available data should be self evident. Only when the human operator model is uniquely defined by data can its parameter values be used to predict the effects of external stresses, plant dynamics, disturbance inputs, motion cues, instrument scanning and other factors.

Having selected an identifiable model structure, the next important task is the choice of a particular parameterization for describing the available data. As mentioned earlier, fit error is not an adequate criterion in selecting the model parameterization. For example, suppose we have N observations on the input $u(t)$, $t = 1,2,\cdots,N$ and the output $x(t)$, $t = 1,2\cdots,N$. This input-output data could have a mixed autoregressive moving average (ARMA) structure as its true model. However, this data may also be modeled by a sufficiently long autoregressive (AR) parameterization. Jenkins and Watts [10] have demonstrated the inefficiency of the direct estimation of the impulse response function as compared to its parametric (difference or differential equation) representation. Replacing a lower order ARMA model with a high order AR representation would lead to similar problems and hence inefficient parameter estimates. For predictive purposes, a model parameterization with fewer parameters is preferable. Therefore, the objective of modeling, including human operator modeling, must be to develop identifiable structures and parsimonious parameterizations. Model predictions could be seriously deficient if frugality is not exercised in model structure formulation and parameterization.

This report reviews the formulation and validation of the SOCM from a structural identifiability and model parameterization viewpoint. The SOCM as well as the standard laboratory data used in its verification are summarized in the following section. Revisions to the SOCM made by BBN [11-12], (hereafter referred to as the Revised Optimal Control Model - ROCM) to account for motion cue effects are discussed in Section 3. Section 4 comments on model validation procedures and on the identifiability of the optimal control model structure. A simple example is used to illustrate some of the ambiguities related to parameter selection or estimation. Implications of considering the SOCM and the ROCM as just specific representations within the class of optimal control models are presented in Section 5. A general set of hypotheses describing the class of optimal control models is defined. Alternate hypotheses within this general set are shown to be capable of describing the standard laboratory tracking data corresponding to k, k/s and k/s^2 plant dynamics.

It is assumed that the reader is familar with the SOCM as proposed by Kleinman, Baron and Levison approximately ten years ago. This preliminary work was accomplished by this author under a 30 day consulting contract (Request No. 22-ME-4) with the University of Dayton Research Institute for the 6570th Aerospace Medical Research Laboratory (AMRL), Wright-Patterson Air Force Base, Ohio. Even so, the results to date are very interesting and indicate that further research into the subject should be rewarding.

II. THE STANDARD OPTIMAL CONTROL MODEL

The development of the standard optimal control model constituted a radical departure from human operator modeling technology prevalent during that period. Quasi-linear describing function models and, to some extent, discrete sampled data models were the only human operator models with any theoretical basis available to the manual control community. The SOCM, therefore, had the burden of duplicating the peculiar human psycho-physical model. These included a time delay, a remnant or noise component, a central nervous system information processing unit referred to as an equalizer element, and a so-called neuromuscular lag reflecting the neuromotor system dynamics. The basic assumption in the SOCM is that the trained human operator in a precision control task behaves like an optimal controller and information processor with respect to the given task objectives. Imposition of the above mentioned psycho-physical constraints on this optimality assumption leads to the well known SOCM structure shown in Figure 1.

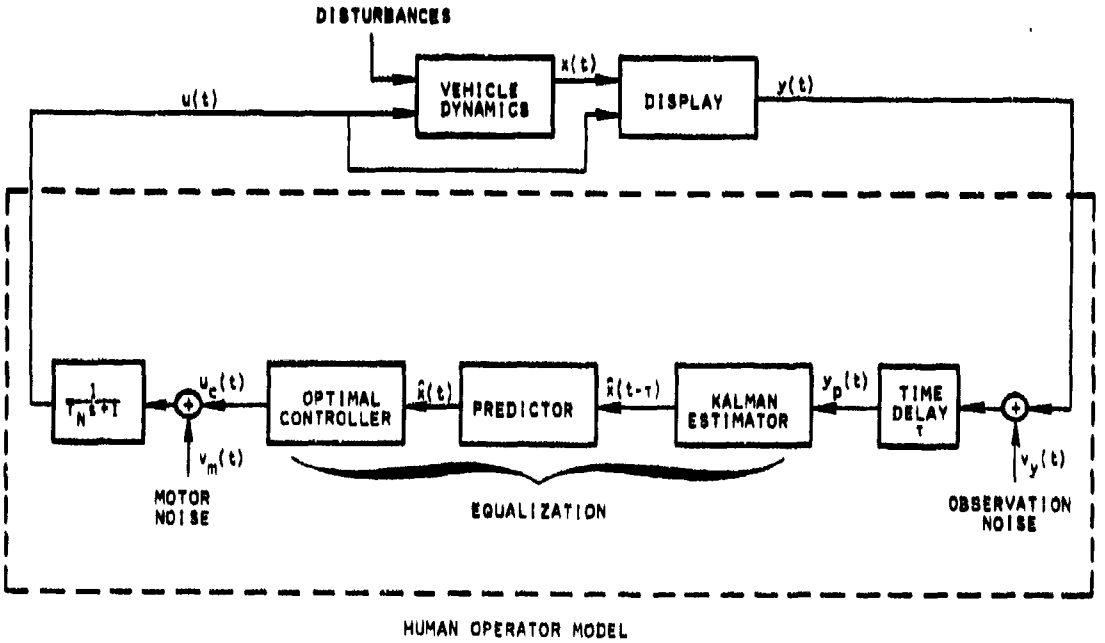


FIGURE 1. CONTROL-THEORETIC MODEL OF HUMAN RESPONSE BEHAVIOR

The SOCM consists of four tandem components: corresponding to models for human perception, information processing, control and neuromuscular dynamics. Specifically, the SOCM assumes that (1) the time delay τ is lumped at the perceptual end and the human operator perceives delayed noisy observations of the displayed variables and their first derivatives, (2) the trained human observer has an internal state variables model for the plant dynamics, commanded inputs and stochastic disturbances that is identical with the actual linearized system model, (3) the operator chooses a control law that minimizes a quadratic cost functional in terms of the state vector, the control rate, and (4) the operator has neuromuscular limitations reflected in a noisy, filtered, implementation of optimal control feedback.

Thus, the SOCM is completely parameterized by the quadratic cost functional weighting matrices, the time delay in the observations and the motor and observation noise covariances. Note that the so-called neuromotor time constant T_N is implicitly characterized by the weighting matrix on the control rate vector. The SOCM is a special case of the state variable model. The model structure and order are completely determined by the models for the plant/noise dynamics and the weighting matrices in the task cost functional. The critical problem in using the SOCM for a specific manual control task lies in the selection of the model parameter values. The basic approach used in the past has been to iterate on the model parameters on a trial-and-error basis so as to get a good fit or match between model outputs and actual measured data. Typically, spectral data on the equivalent input-output describing function magnitude and phase characteristics, and the remnant, obtained from standard laboratory manual tracking task experiments has been used. A review of this standard laboratory experimental data is given in the following because of its relative historic importance in validating the SOCM and the ROCM.

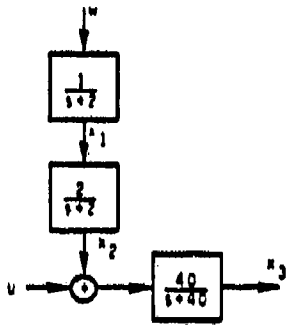
2.1 Review of Standard Laboratory Experiments/Data

Experimental data from three single axis manual tracking tasks corresponding to position control or k dynamics [5], rate control or k/s dynamics [6] and acceleration control or k/s^2 dynamics [5] has been extensively used as a baseline data for determining or selecting human operator model parameters. A signal consisting of a sum of 13 sine waves was used to generate the disturbance input. The input spectrum approximated a first order noise with a bandwidth of 2 rad/s for the k/s and k/s^2 experiments and a second order noise process with a bandwidth of 2 rad/s for the k experiments. The disturbance input was added in parallel to the human operator's control input for the k and k/s tasks but was added to the plant output rate for the k/s^2 experiments.

Internal models for the plant/noise dynamics and other elements used in the SOCM are given below. The only approximation made is in the representation of the sum of sine waves disturbance input as a state variable noise model.

k Dynamics:

Block Diagram:



Note: $\frac{40}{s + 40}$ introduced for analytical reasons.

Plant Dynamics: $\dot{x}_3 = -40x_3 + 40\,(x_2 + u)$

Disturbance Model: $\dot{x}_1 = -2x_1 + w \;;\; W = \sigma_w^2 = 5.425$

$\dot{x}_2 = -2x_2 + 2x_1$

Perception Model: $y_{p1} = y_1(t - \tau)$,

$y_{p2} = y_2(t - \tau)$

where $y_1 = x_3 + v_{y2}$

$y_2 = \dot{x}_3 + v_{y2}$

τ = time delay

Quasratic Cost: $J = \int [x_3^2 + g\,\dot{u}^2]\,dt$

Motor Model: $T_n\dot{u} + u = u_c + v_m$

SOCM parameters: Control rate weighting g
Observation noise covariances V_{y1}, V_{y2}
Motor noise covariance V_m

where $V_{y1} = \rho_{y1}\,\pi\overline{y_1^2}$ ($\rho_{y1} = 0.01 \equiv -20$ dB)

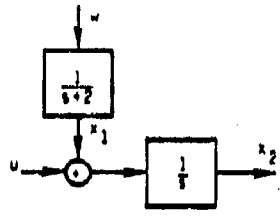
$V_{y2} = \rho_{y2}\,\pi\overline{y_2^2}$ ($\rho_{y2} = 0.01 \equiv -20$ dB)

$V_m = \rho_m\,\pi\overline{u_c^2}$ ($\rho_m = 0.003 \equiv -25$ dB)

Figure 2. shows the k dynamics describing function (magnitude and phase) and remnant (injected at plant output) data as well as the SOCM output fit for g = 0.01, τ = 0.15, and equal attention allocation to y_1 and y_2. SOCM output seems to match experimental data farily well except for the phase discrepancy at low frequencies.

k/s Dynamics:

Block Diagram:



Plant Dynamics: $\dot{x}_2 = x_1 + u$

Disturbance Model: $\dot{x}_1 = -2x_1 + w$

$W = \sigma_w^2 = 8.8$

Perception Model: $y_{p1} = y_1(t - \tau)$

$y_{p2} = y_2(t - \tau)$

$y_1 = x_2 + v_{y1}$

$y_2 = \dot{x}_2 + v_{y2}$

τ = time delay

Quadratic Cost: $J = \int [x_2^2 + g\,\dot{u}^2]\,dt$

Motor Model: $T_n\dot{u} + u = u_c + v_m$

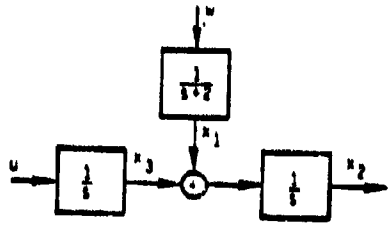
SOCM Parameters: Control rate weighting g
Observation noise covariances V_{y1}, V_{y2}
Motor noise covariance V_m

where V_{y1}, V_{y2} and V_m are defined as in k example above.

Figure 3 shows the k/s dynamics describing function and remnant data along with the SOCM output match for g = 0.00017, τ = 0.15 and equal attention to y_1 and y_2. As with the k data, SOCM outputs seem to match the experimental data fairly well except for the phase error at low frequencies.

k/s^2 Dynamics:

Block Diagram:



Plant Dynamics: $\dot{x}_2 = x_1 + x_3$

$\dot{x}_3 = u$

Disturbance Model: $\dot{x}_1 = -2x_1 + w$

$W = \sigma_w^2 = 0.217$

Perception Model: $y_{p1} = y_1(t - \tau)$

$y_{p2} = y_2(t - \tau)$

$y_1 = x_2 + v_{y1}$

$y_2 = \dot{x}_2 + v_{y2}$

τ = time delay

Quadratic Cost: $J = \int [x_2^2 + g\,\dot{u}^2]\,dt$

Motor Model: $T_n u + \dot{u} = u_c + v_m$

SCOM parameters: g, V_{y1}, V_{y2} and V_m as before.

Figure 4 shows the human operator describing function and remnant data for the k/s^2 experiments. SOCM outputs are also shown on the same figure for $g = 0.00007$, $\tau = 0.21$ and equal attention allocation to y_1 and y_2. Model output seems to match data over the complete measurement frequency range.

In each of the above cases, the equivalent describing function for the human operator is given by

$$u(s) = h_1(s)\, y_1(s) + h_2(s)\, y_2(s)$$

$$= [h_1(s) + s\, h_2(s)]\, y_1(s)$$

$$\text{or } \left[\frac{u}{y_1}(s)\right]_e = h_e(s) = h_1(s) + s\, h_2(s)$$

The equivalent remnant spectrum injected at the plant output is given by

$$\Phi_{y_1 y_1}(s) = \frac{|h_1(s)|^2 V_{y1} + |h_2(s)|^2 V_{y2}}{|h_e(s)|^2} + \frac{V_m}{|T_n s + 1|^2\, |h_e(s)|^2}$$

similarly, the equivalent remnant spectrum injected at the plant output rate is given as

$$\Phi_{y_2y_2}(s) = s^2 \Phi_{y_1y_1}(s)$$

$$= \frac{|s\, h_1(s)|^2 V_{y_1} + |s\, h_2(s)|^2 V_{y_2}}{|h_e(s)|^2} + \frac{|s|^2 V_m}{|T_n s+1|^2 |h_e(s)|^2}$$

The theoretical model based mean squared values of plant output, plant output rate and control are compared against measured data in Table I below

TABLE I

MEASURED VERSUS THEORETICAL CLOSED-LOOP PERFORMANCE

System	Mean squared plant output		Mean squared plant output rate		Mean squared control	
	Meas.	Theor.	Meas.	Theor.	Meas.	Theor.
k	0.13	0.15	4.8	10.725+	0.53	0.486
k/s	0.13	0.114	3.1	2.97	4.2	3.76
k/s^2	0.014	0.0133	0.10	0.102	1.43	1.15

Note the discrepancy in the mean squared values of plant output rate for the k system. The theoretical value of 10.725 is more than twice the measured value. This number in the report by Kleinman and Baron [4, Table 1/p37] is given as 5.3 which is closer to the measured value of 4.8 and has a footnote attached to it explaining that the number given is over the measurement frequency range $\omega < 32$ rad/s. Clearly, if exceptions are to be made they should be applied uniformly across the data and not to a selected set of variables. Also note that the value of W for the k system is not given explicitly in any of the previous reports and had to be determined iteratively using the PIREP [13] program resident at NASA-Ames Research Center.

The results shown earlier demonstrate good agreement between the SOCM model and the measured data for k, k/s and k/s^2 manual control tasks. The SOCM model parameters - ρ_{y1}, ρ_{y2}, ρ_m, g and τ are selected after considerable number of iterations such that the model describing function (magnitude and phase) and remnant match measured spectral estimates. It is, therefore, inaccurate to refer to the good model match as a test of the models predictability. Model parameters were not fixed apriori at the values which provide a good match to measured data. On the contrary, model parameters were determined iteratively, so as to match measured describing function and remnant characteristics.

It must also be emphasized that the standard laboratory data for k, k/s and k/s^2 dynamics discussed above is not as reliable as one would expect, considering that this data is referred to as the baseline for validating any human operator model and in particular, revisions or modifications to the SOCM [19]. For example, the k and k/s^2 experimental data was obtained in support of the XV-5A hovering task application with the objective of determining the various observation noise ratios [5]. The measured describing function and remnant reflect the average of two runs per subject. In addition the disturbance input power levels were not fixed throughout the experiments. As stated in reference 5, (p57) "different levels of ability among subjects necessitated the use of different input levels to achieve the same signal (plant output) variance". In addition describing function and, in particular, the remnant data shown in more recent references (e.g., [3,4] is not identical to the data in the original sources [5,6] from which they were derived.

The above discussion shows that the k and k/s^2 system data has its drawbacks which should be considered in evaluating the relative merits of one model against another. The k/s data is relatively better and includes an average of 4 subjects with 4 trials per subject.

III. THE REVISED OPTIMAL CONTROL MODEL

Section 2 shows that the SOCM is able to describe almost all of the key features of the describing function and remnant data for the standard laboratory k, k/s and k/s^2 experiments. The primary deficiency evident from Figures 2-4 is the inability of the model to duplicate the, so called, low frequency droop in the phase characteristics for the k (Fig. 2) and k/s (Fig. 3) compensatory regulation tasks. In other words, the SOCM is unsuccessful in reproducing the $(\exp(\alpha/s))$ low frequency phase droop term in the Extended Crossover Describing Function Model (second approximation) [14] for the human operator. Low frequency phase lags have been ignored in SOCM model matching procedures because they have no substantial effect on closed loop man-vehicle performance. This is the result of the relatively high system loop gain ($h_e h_c(s) \gg 1$) at the low frequencies where the phase droop is evident. However, it has been shown that low frequency phase droop in the equivalent input-output describing function is eliminated in closed-loop "pursuit" tracking tasks where the human operator has access to either the tracking disturbance input (in visual pursuit tasks) [15] or the plant output (in motion cue tasks on a simulator) [11] in addition to the usual tracking error in compensatory tracking tasks. Consequently, it is important that a human operator model describe low frequency phase characteristics of the equivalent input-output describing function if one is to understand the underlying mechanisms in human sensory feedback cue utilization and adaptation.

The Revised Optimal Control Model (ROCM), as developed by Bolt, Beranek and Newman Inc. (BBN) [11,12], represents a refinement of the SOCM incorporating modifications that give it the ability to describe the low frequency phase characteristics. The SOCM has a peculiar formulation of the control law that introduces a so called neuromotor system structure as part of the equivalent input-output describing function. To be specific, the SOCM

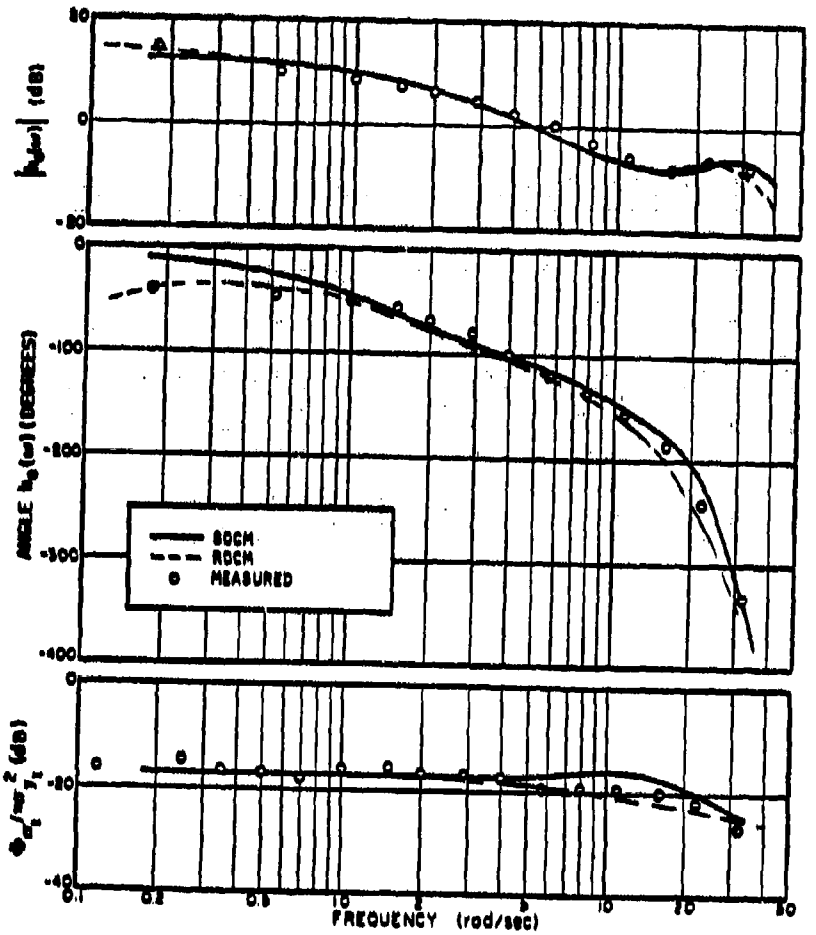


FIGURE 2. MEASURED AND PREDICTED FREQUENCY DOMAIN MEASURES, k DYNAMICS (Average of three subjects)

assumes a quadratic cost functional J in state vector x and control rate $\dot{u}$ of the general form (for scalar u)

$$J = \int (x^T Q x + g\,\dot{u}^2)\,dt \qquad (3.1)$$

where

$$\dot{x} = Fx + Gu + \Gamma w \qquad (3.2a)$$

$$\dot{u} = \tilde{u}. \qquad (3.2b)$$

The optimal control law is obtained by solving the appropriate Riccati equation as

$$\tilde{u} = \dot{u} = -\lambda^T \hat{x} - \lambda_u u \qquad (3.3)$$

λ = optimal gain feedback on the plant state estimate $\hat{x}$

λ_u = optimal gain feedback on control u

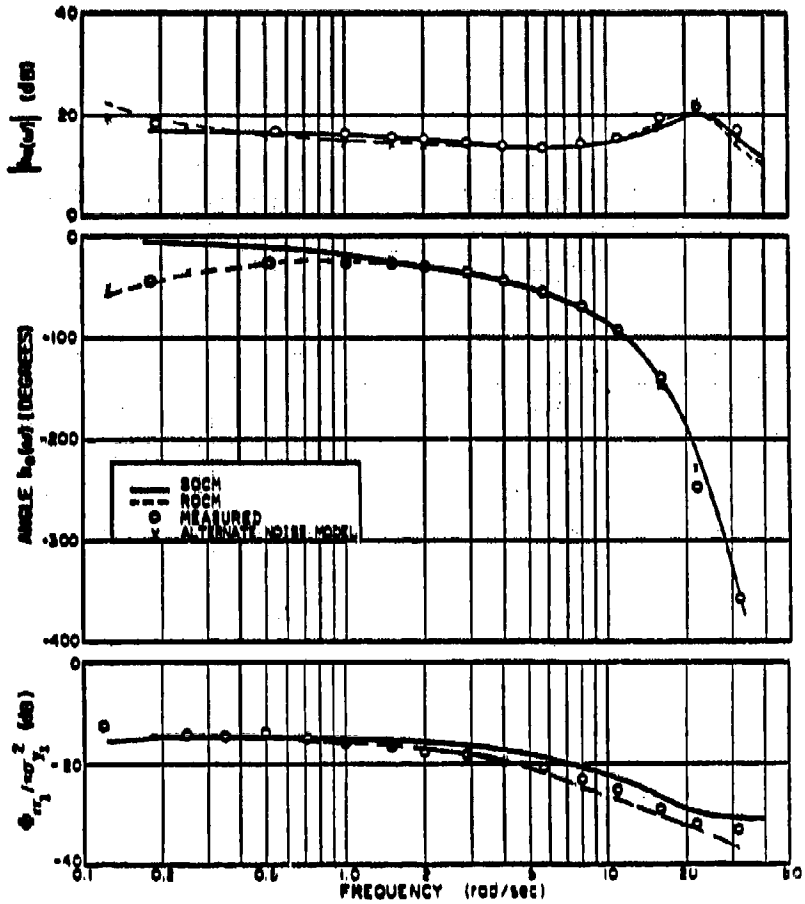


FIGURE 3. MEASURED AND PREDICTED FREQUENCY DOMAIN MEASURES, k/s DYNAMICS (Average of four subjects)

Then $\dot{u} + \lambda_u u = -\lambda^T \hat{x}$

or $$T_N \dot{u} + u \equiv \frac{1}{\lambda_u}\dot{u} + u = -\frac{1}{\lambda_u}\lambda^T\hat{x} = -\ell\hat{x} = u_c \qquad (3.4)$$

Thus, $$T_N \equiv \frac{1}{\lambda_u} \text{ and } \ell \equiv -\frac{\lambda^T}{\lambda_u} \qquad (3.5)$$

The SOCM equivalent describing function is given as

$$u(s) = \left(\frac{1}{T_N s+1}\right)\left[\frac{u_c}{y}(s)\right] y(s) = \frac{-\ell}{(T_N s+1)}\left[\frac{\ell}{y}(s)\right] y(s) \qquad (3.6)$$

where display vector

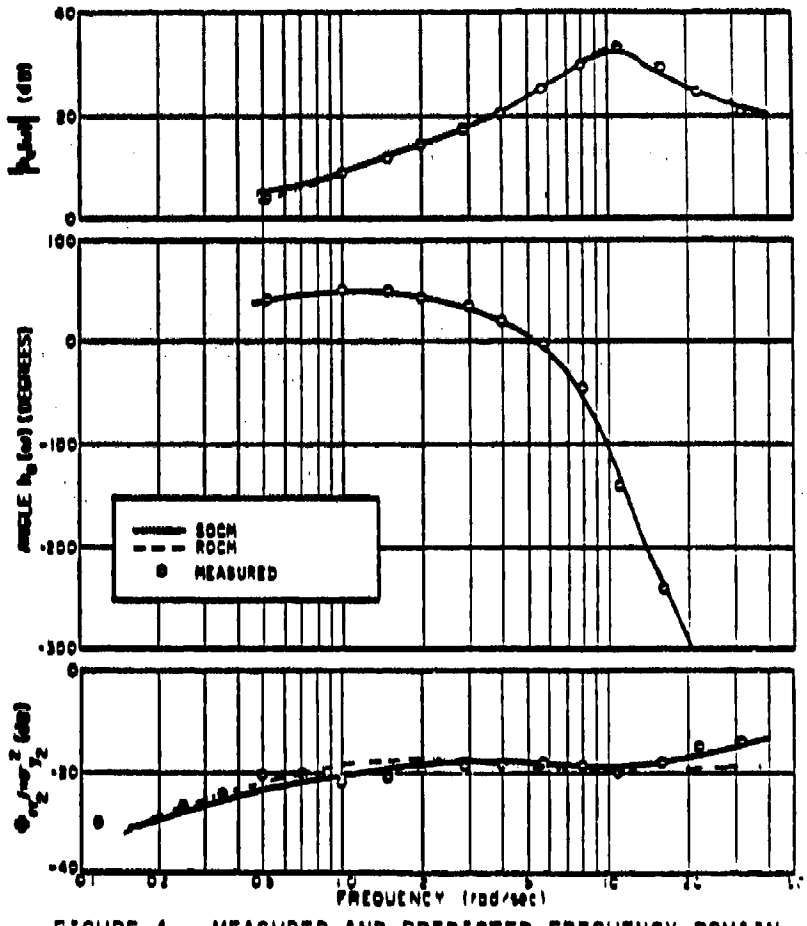


FIGURE 4. MEASURED AND PREDICTED FREQUENCY DOMAIN MEASURES k/s^2 DYNAMICS (Average of three subjects)

$$y = Hx + Du + v_y \tag{3.7}$$

and perceived vector

$$y_p = y(t-\tau) \tag{3.8}$$

τ = delay time; v_y = observation noise vector $\hat{x}(t)$ = minimum mean squared error (mmse) estimate of $x(t)$ given $y_p(t)$ (3.9)

Note that a motor noise v_m is added to u_c in computing the closed loop system performance and the Kalman filter for $\hat{x}$. Thus

$$T_N\dot{u} + u = u_c + v_m \tag{3.10}$$

The Kalman filter utilizes equations (3.2a and (3.10) as the plant state equations and equations (3.7) and (3.8) for the measurement vector, to give $\hat{x}$ and $\hat{u}$. However, the control is computed using

$$u_c = -[\ell, 0]\begin{bmatrix}\hat{x}\\ \hat{u}\end{bmatrix} \tag{3.11}$$

where the control gain vector $[\ell, 0]$ is not the optimal gain vector for the system equations (3.2a) and 3.10) but instead is optimal for the system given by the equations (3.2a) and (3.2b). This fact implies that the resulting SOCM is not "optimal" in the linear, quadratic, gaussian sense as the name suggests. This suboptimality of the control law is the result of using equation (3.10) instead of (3.2b) in an attempt to incorporate a neuromuscular lag $(1/T_N\ s+1)$ into the human operator model. It is easily shown that the so-called neuromuscular lag $(1/T_N\ s+1)$ is nearly cancelled by the Kalman filter zero $[T_N's+1]$, $T_N' \cong T_N$, for nominal values of the motor noise covariance V_m.

Clearly, the correct procedure would be to use

$$\dot{u} = \tilde{u}_c + v_u \tag{3.12}$$

where the motor noise v_u is treated as a disturbance to control rate

$$\tilde{u}_c = -(\lambda^T, \lambda_u)\begin{pmatrix}\hat{x}\\ \hat{u}\end{pmatrix} \tag{3.13}$$

where $\hat{x}$ and $\hat{u}$ are obtained from the Kalman filter utilizing equations (3.2a), (3.12) for the plant state equations and (3.7) and (3.8) as the measurement equations. The resulting controller is optimal in the LQG sense if all the assumptions are satisfied. Note that

$$u(s) = \frac{1}{s}\left[\frac{\tilde{u}_c}{y}(s)\right]\ y(s) \tag{3.14}$$

It is easily shown that $(\tilde{u}_c/y)$ has a zero $(s + \varepsilon)$ where $\varepsilon \equiv 0$ for $V_u \equiv 0$ [16]. $\varepsilon \approx 0$ for nominal values of V_u. This implies

$$u(s) = \left(\frac{s+\varepsilon}{s}\right)\left(\frac{N(s)}{D(s)}\right) y(s) \tag{3.15}$$

where the $(s + \varepsilon)/s$ provides the low frequency droop observed in compensatory tracking describing function data. This idea has been used by BBN to modify the SOCM. The ROCM assumes the above formulation with one major modification: Two types of motor noise are proposed. A so called "pseudo" motor noise v_u

(cov $v_u = V_u$) that is used in the internal model for Kalman filter estimation and a so called "driving" motor noise v_m (cov $v_m = V_m$) which is a disturbance in the commanded control rate $\tilde{u}_c$ and drives the closed loop man-vehicle system. In short, use

$$\hat{u} = \tilde{u}_c + v_u \qquad \text{cov}\,(v_u) = V_u \tag{3.16}$$

for derivation of the Kalman filter equation and

$$\hat{u} = \tilde{u}_c + v_m \qquad \text{cov}\,(v_m) = V_m \tag{3.17}$$

in computing the mean square values of system variables and in defining the equivalent injected remnant. In general $V_u \neq V_m$ since it is difficult to match both describing function and remnant data with $V_u = V_m$ [11].

A block diagram of the ROCM is shown in Figure 5. The pseudo motor noise v_u is shown dashed to indicate that it does not drive the system but is merely used in the human's internal model formulation of the plant/noise dynamics. ROCM parameters were adjusted on a trial and error basis to match the k, k/s and k/s^2 describing function and remnant data in Figures 2-4. The specific ROCM parameters used are shown in Table II and the corresponding model outputs are given by dashed lines in Figures 2-4. Note that both the SOCM and ROCM show insignificant differences in the mid and high frequency regions. The ROCM, however, provides the low frequency droop match to data which the SOCM does not. Both the ROCM and SOCM have approximately the same values for the

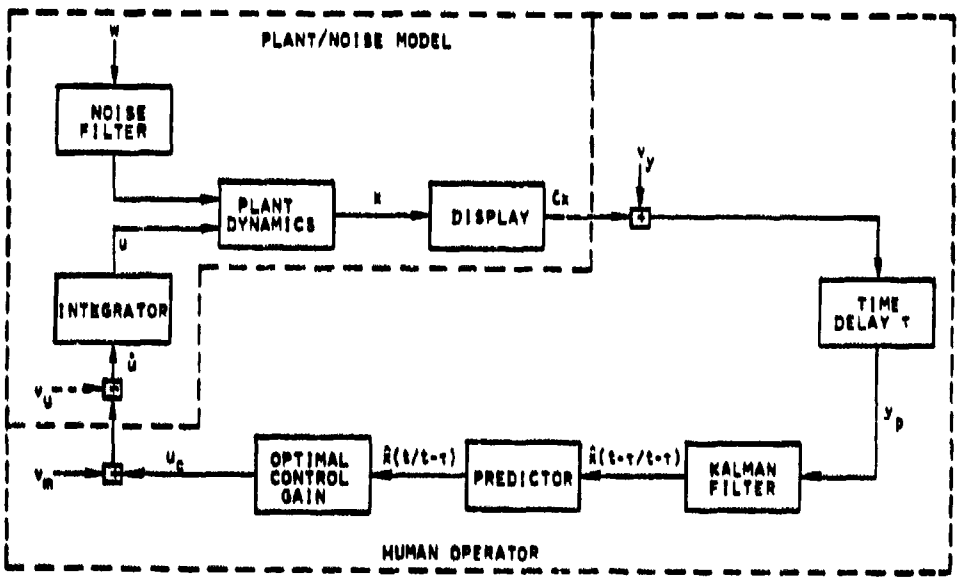


FIGURE 5. STRUCTURE OF THE REVISED OPTIMAL CONTROL MODEL (ROCM) FOR THE HUMAN OPERATOR

TABLE II SOCM/ROCM PARAMETERS
-k, k/s, k/s^2 DATA-

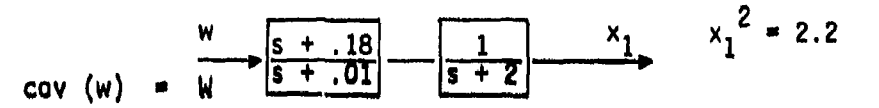
MODEL PARAMETER	k		k/s		k/s^2	
	SOCM	ROCM	SOCM	ROCM	SOCM	ROCM
CONTROL RATE WEIGHTING g	0.01†	0.01	0.00017	0.0003†	0.00007	0.00007
MOTOR TIME CONSTANT* T_N	0.1112†	0.1112	0.0807	0.093	0.1016	0.101
TIME DELAY τ	0.18	0.18	0.18	0.17	0.21	0.21
PSEUDO-MOTOR NOISE/SIGNAL RATIO (dB) P_u	—	-20	—	-21	—	-40
DRIVING MOTOR NOISE/SIGNAL RATIO (dB) P_m	-25	-∞	-25	-∞	-25	-∞
BASE OBSERVATION NOISE/SIGNAL RATIO (dB) P_y	-20	-20	-20	-21	-20	-20
ATTENTION TO PLANT OUTPUT f_{y1}	1.0	1.0	1.0	1.0	1.0	0.1
ATTENTION TO PLANT OUTPUT RATE f_{y2}	1.0	0.5	1.0	1.0	1.0	1.0
DISTURBANCE INPUT PROCESS NOISE VARIANCE W	5.425†	5.425	5.8	5.8	0.217	0.217

†OBTAINED THROUGH ITERATIVE USE OF PIREP PROGRAM
*FOR k/s DYNAMICS: $T_N = 0.707\ g^{1/4}$

control rate weighting g and the observation noise/signal ratios (ρ_y). However, the latter is misleading because the fractional allocation of attention to plant output f_{y1} and rate f_{y2} is not held constant and equal to one in the ROCM except for the k/s task. In fact (f_{y1}, f_{y2}) is (1.0, 0.5) for the k task and (0.1, 1.0) for the k/s^2 task. The driving motor noise v_m is set to zero for all three cases. The pseudo-motor noise/signal ratio (ρ_m) is set equal to the base observation noise/signal ratio ρ_y for the k and k/s tasks but is much smaller (-40 dB) for the k/s^2 task. This is obviously because the SOCM is able to duplicate the k/s^2 describing function and remnant data.

The ROCM represents one approach to modifying the SOCM assumptions in order to describe existing laboratory data over the complete range of the input frequencies (.18 - 32.0 rad/s). As pointed out by BBN [11], the pseudo motor noise introduced in the ROCM does not seem to have any particular rationale that can be attributed to human psycho-physical or physiological characteristics. Pseudo motor noise must merely be treated as a mechanism or artifice that contributes to the generation of the low frequency phase lag or droop in the k, k/s and k/s^2 regulation tasks. Therefore it is reasonable to believe that there may be one or more alternate modifications to the SOCM assumptions that may be capable of reproducing the low frequency phase characteristics in the human operator describing function and other data over a wide frequency range. One such alternate hypothesis is to assume a different model

for the sum of sine waves disturbance input from that used in the SOCM. A preliminary investigation of the k/s data shows that a sum of sine waves model given below



$$\text{cov}(w) = W = 3.375$$

where $x_1^2 = 2.2$ is the mean squared value of the disturbance input, gives rise to a low-frequency droop. Figure 3 shows that the magnitude and phase of the SOCM with this modified internal model for the disturbance input match the available data as well as the ROCM fit. Remnant output was not computed in this preliminary investigation, but should match available data. The only modification made here is in the internal model for the disturbance input by adding a lag-lead filter to the existing first order filter at 2 rad/s. The remaining parameters of the SOCM remain unchanged. Both the ROCM and the SOCM/with modified disturbance model reflect alternate ways of representing plant and/or disturbance uncertainty. This lack of uniqueness of model parameterization in describing a given set of data can cause problems in certain applications. This subject is discussed in the following section.

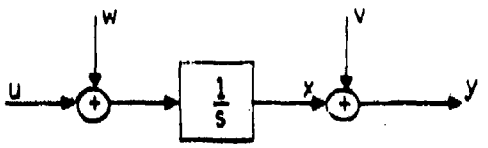
IV. MODEL VALIDATION PROCEDURES

The usual approach to determining the parameters of an optimal control model for the human operator is to iterate on the model parameters until model responses or outputs reasonably match measured data. The procedure is heuristic and its rate of convergence to an acceptable parameter set depends intimately upon the skill and experience of the investigator. However, there is no guarantee that the selected parameters uniquely characterize the measured human response data. In other words, it is quite possible to have two or more parameter sets describe the same experimental data. In addition, differences can arise in the variances obtained on the parameter estimates. The non-uniqueness of the parameter estimates is related to the identifiability properties of the model structure while variances in parameter estimates are affected more by the parsimoniousness of the chosen parameterization.

Results of the previous section demonstrate the feasibility of having two or more parameterizations for a given set of data. This is evident by the fact that both the ROCM and the SOCM / with modified disturbance model are equally able to describe k/s describing function data over the frequency range of interest. This gives additional justification for investigating alternate optimal control model structures for the human operator. However, before proceeding with that a simple optimal control model example is developed to illustrate the concept of structural identifiability.

Identifiability Example

Consider a regulation task as shown below



Plant/Noise Model: $\dot{x} = u + w$ $\quad cov(w) = W$

$y = x + v$ $\quad cov(v) = V$

Quadratic Cost: $J = \int (x^2 + gu^2)\, dt$

Kalman filter: $\dot{\hat{x}} = u + k(y - \hat{x})$

Optimal Control: $u = -\lambda\hat{x}$, $\lambda = \frac{1}{\sqrt{g}}$

Kalman Gain: $k = \sqrt{W/V}$

Equivalent Describing Function: $$\frac{u}{y}(s) = \frac{-\frac{1}{\sqrt{g}}\sqrt{\frac{W}{V}}}{s + \frac{1}{\sqrt{g}} + \sqrt{\frac{W}{V}}}$$

Equivalent Remnant: v

This example assumes zero time delay and no motor noise. The measured data includes u and y. Hence, the describing function parameters and remnant can be identified uniquely from data. The identifiable parameters are V, and

$$\frac{1}{\sqrt{g}}\sqrt{\frac{W}{V}} = d$$

$$\frac{1}{\sqrt{g}} + \sqrt{\frac{W}{V}} = c$$

The unknowns are $\frac{1}{\sqrt{g}}$, $\sqrt{\frac{W}{V}}$, V

If W is exactly known to the human operator then g and V can be identified uniquely. However, if W is not known, then there exist two sets of parameters with the same describing function and remnant.

Let $a = \frac{1}{\sqrt{g}}$, $b = \sqrt{\frac{W}{V}}$, then

$$a + b = c \quad \text{and} \quad ab = d$$

$$(a - b)^2 = (a + b)^2 - 4ab = c^2 - 4d$$

$$a - b = \pm \sqrt{c^2 - 4d}$$

$$a = \tfrac{1}{2}\left[c \pm \sqrt{c^2 - 4d} \right]$$

$$b = \tfrac{1}{2}\left[c \mp \sqrt{c^2 - 4d} \right]$$

The two solutions are $\left\{ \frac{1}{2}c + \frac{1}{2}\sqrt{c^2 - 4d},\ \frac{1}{2}c - \frac{1}{2}\sqrt{c^2 - 4d},\ V \right\}$

and $\left\{ \frac{1}{2}c - \frac{1}{2}\sqrt{c^2 - 4d},\ \frac{1}{2}c + \frac{1}{2}\sqrt{c^2 - 4d},\ V \right\}$ respectively

If $c = 3$ and $d = 2$

$$a = \frac{1}{\sqrt{g}} = 2 \text{ or } 1 \quad \text{and} \quad b = \sqrt{\frac{W}{V}} = 1 \text{ or } 2 \text{ respectively}$$

The two solutions are

$$\left\{ \frac{1}{\sqrt{g}}, \sqrt{\frac{W}{V}}, V \right\} = \left\{ 2, 1, V \right\} \text{ or } \left\{ 1, 2, V \right\}$$

This simple example demonstrates the concept of non-unique parameterizations in context of the optimal control model structure. Unfortunately, explicit solutions of the Riccati equations for realistic higher order models is not possible; (except for pathologic cases) consequently, analytic solution of multiple parameterizations for a given set of data cannot be provided.

The simple example considered here had two distinct parameter sets capable of describing available data. It is conceivable that an infinite number of parameter sets could exist for more complex models. An example would be a transfer function and remnant given by

$$\frac{u}{y}(s) = \frac{1}{s + p + q} = \frac{1}{s + c}$$

remnant = v

where $p + q = c$ = constant corresponds to an infinite number of solutions (p, q) on the line $p + q = c$. Numerical optimization techinques for parameter estimation would exhibit extremely high variances on parameters p and q. The above discussion underscores the importance of understanding the structural identifiability properties of a given model structure. However, identifiability is not the only property which must be analyzed during model development. Another property of the models is that of parsimony or degree of parameterization. An identifiable model that is over-parameterized can also lead to difficulty in parameter estimation if explicit parameter identification techniques are utilized for the purpose. However, if the parameters are determined by the heuristic iteration process with the human investigator making judgements on model matching accuracy, then a solution can be arrived at very easily. In fact, the quality of the data match model is not adequate for predictive purposes. Unfortunately, the word "prediction" is too often abused in the literature. For example, SOCM model describing functions and remnant in k, k/s and k/s^2 tasks are referred to as "predicted frequency domain measures" when in fact they were obtained after trial and error iteration on the parameters to fit measured data.

V. ROCM IMPLICATIONS AND ALTERNATE OPTIMAL CONTROL MODEL STRUCTURES

Identification of the optimal control model poses some fundamental questions regarding structural identifiability and parsimony in parameterization. Analytical verification of SOCM structural identifiability is a non-trivial task although general results concerning the problem have been presented in the literature [19]. The problem of overparameterization is less complex and techniques exist for determining the subspace of parameter redundancy [20]. The development of the ROCM complicates the overparameterization problem by allowing additional free parameters to the existing SOCM parameter set; namely: (i) free allocation (unequal) of attention levels (f_{y1}, f_{y2}) to position and rate signals, (ii) introduction of two motor noise's-pseudo and driving, and (iii) loss of the constraint on g to give a given value of T_N, the neuromuscular time constant.

The neuromotor system rationale for selection of $\dot{u}^2$ term in the cost functional is no longer applicable in the ROCM. Rather, the term is required to generate the low frequency lag/lead component, $\left(\frac{s + c}{s}\right)$, in the human describing function.

The ROCM is not the only modification to the SOCM that is capable of replicating the low frequency phase droop in the human operator describing function. As shown in Section 3, the SOCM with a modified disturbance model also reproduces the low frequency desbribing function phase characteristics. Thus, one or more model formulations may be capable of describing all of the available experimental data. The ROCM introduced the concept of having an internal model that is different than the actual representation and also eliminated some of the invariance properties of the SOCM across varying task configurations as evidenced by the independence of the fractional attentional allocation to the displayed variable and its first derivative. Clearly, this opens

up possibilities of having pseudo and driving observation noises, the former to be used in state estimation and the later in driving the system response. However, this approach sacrifices the most attractive aspect about the SOCM, namely the relative invariance of some of the parameter selection rules across a wide range of manual control tasks.

The SOCM and ROCM and any modifications to these have one element in common. They represent specific structures within the general class of optimal control human operator models. The SOCM incorporates a fixed set of assumptions in its formulation, an important one being that the human perceives the first derivative or rate of a displayed variable in addition to the displayed quantity. This assumption is not necessary to describe existing laboratory data. An alternate assumption is to assume that the human only perceives the displayed variable y with some observation noise v_y(cov $v_y = V_y$). The remaining optimal control model parameters were iterated until a "good" fit to k, k/s and k/s^2 data was obtained as shown in Figures 6-8 respectively. The resultant alternate optimal control model parameters are shown in Table III.

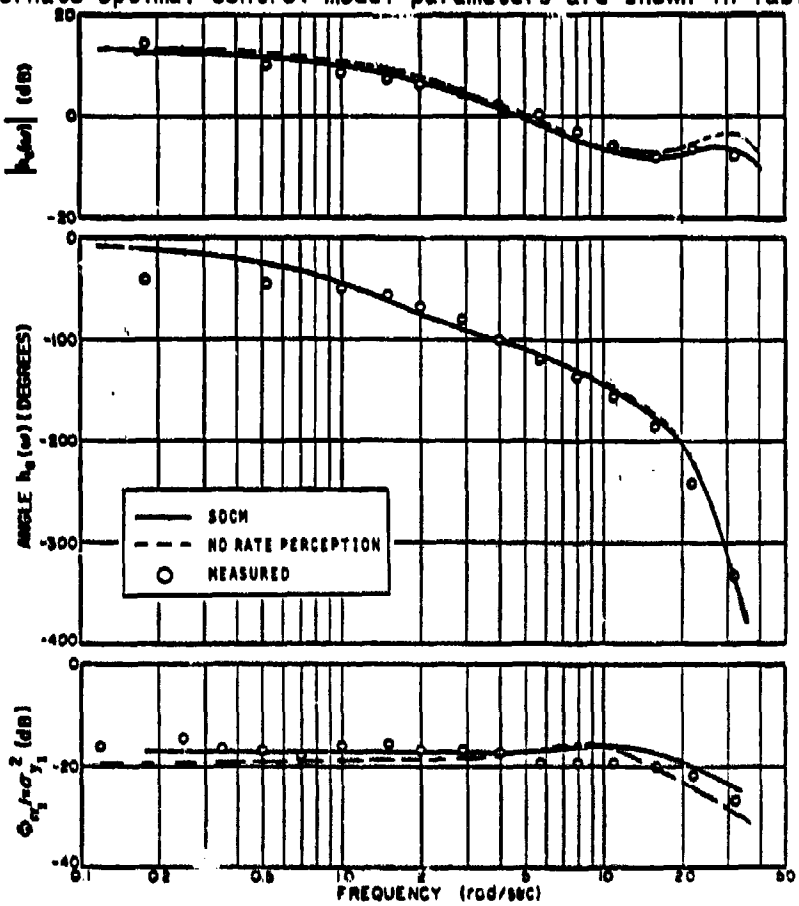


Figure 6. MEASURED AND PREDICTED FREQUENCY DOMAIN MEASURES, k DYNAMICS (Average of three subjects)

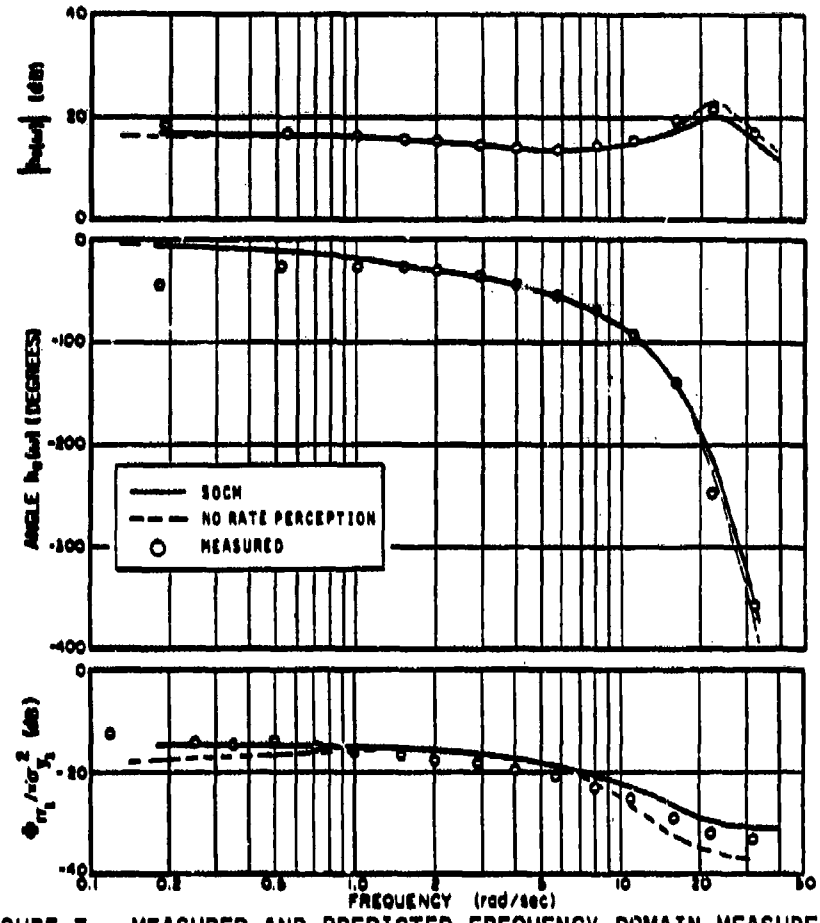


FIGURE 7. MEASURED AND PREDICTED FREQUENCY DOMAIN MEASURES, k/s DYNAMICS (Average of four subjects)

Comparison of these parameters with those for the SOCM in Table II show that the principal differences are in the values of the observation noise/signal ratio ρ_v for the three tasks. Differences in value of g though apparently significant have relatively little impact on the resulting values for T_N (Note $T_N = 0.707\ g^{1/4}$) and the overall crossover frequency. In order to fit the remnant accurately a colored driving motor noise obtained by passing white noise with covariance V_{MD} though a first order filter must be added to the control input u. This motor noise filter $(1/T_M s + 1)$ has a time constant that varies with the plant dynamics; T_M = 0.67, 0.25 and 0.1 seconds and driving motor noise/signal ratio, ρ_{MD} = -12, -16 and -30 dB for k. k/s and k/s^2 dynamics respectively.

Another modification that appears promising is in the formulation of the cost functional. The SOCM uses a quadratic cost functional J in the output

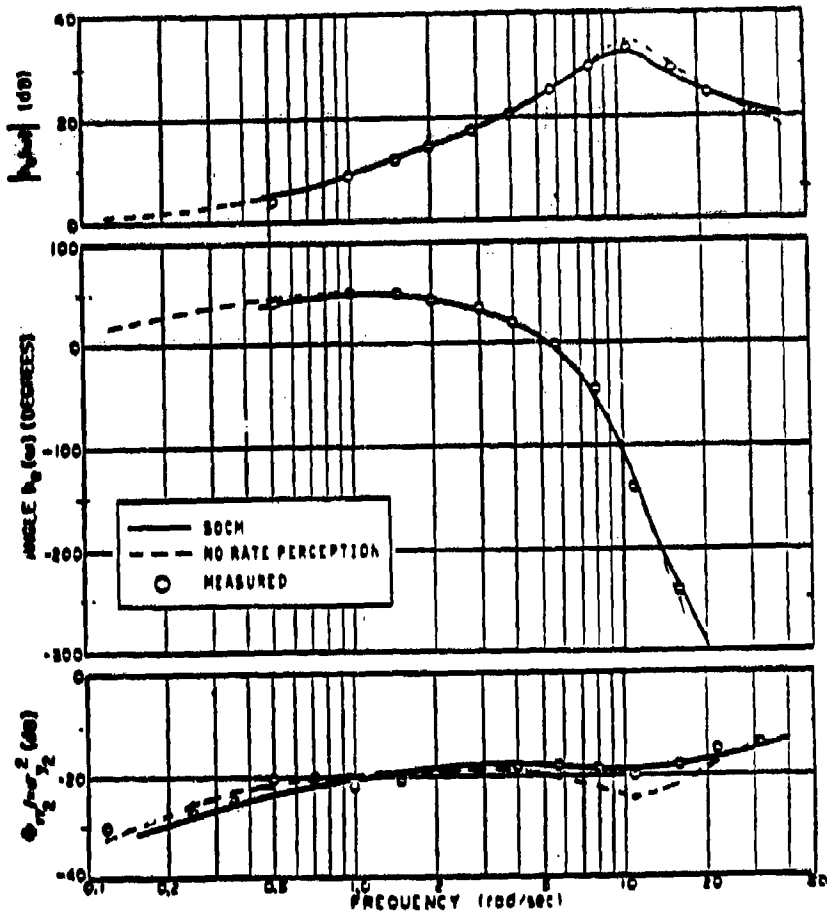


FIGURE 8. MEASURED AND PREDICTED FREQUENCY DOMAIN MEASURES k/s^2 DYNAMICS (Average of three subjects)

vector y and control rate $\dot{u}$. For the k, k/s and k/s^2 dynamics, the following cost expressions are considered (refer to section 2.1 for nomenclature):

$$\text{k Dynamics:} \qquad J = [x_3^2 + g(\ddot{u})^2]\,dt$$

$$\text{k/s Dynamics:} \qquad J = [x_2^2 + g(\dot{u})^2]\,dt$$

$$\text{k/s}^2\text{ Dynamics:} \qquad J = [x_2^2 + g(u)^2]\,dt$$

The basic idea is to choose the order of the u derivative that results in an equivalent type 2 loop transfer function (i.e. $h_e h_c(s) \simeq \frac{1}{s^2}\,\frac{N(s)}{D(s)}$). The type 2 system is desirable as it guaranties zero steady state error to step and ramp disturbance inputs. Figure 9 shows the describing function and remnant

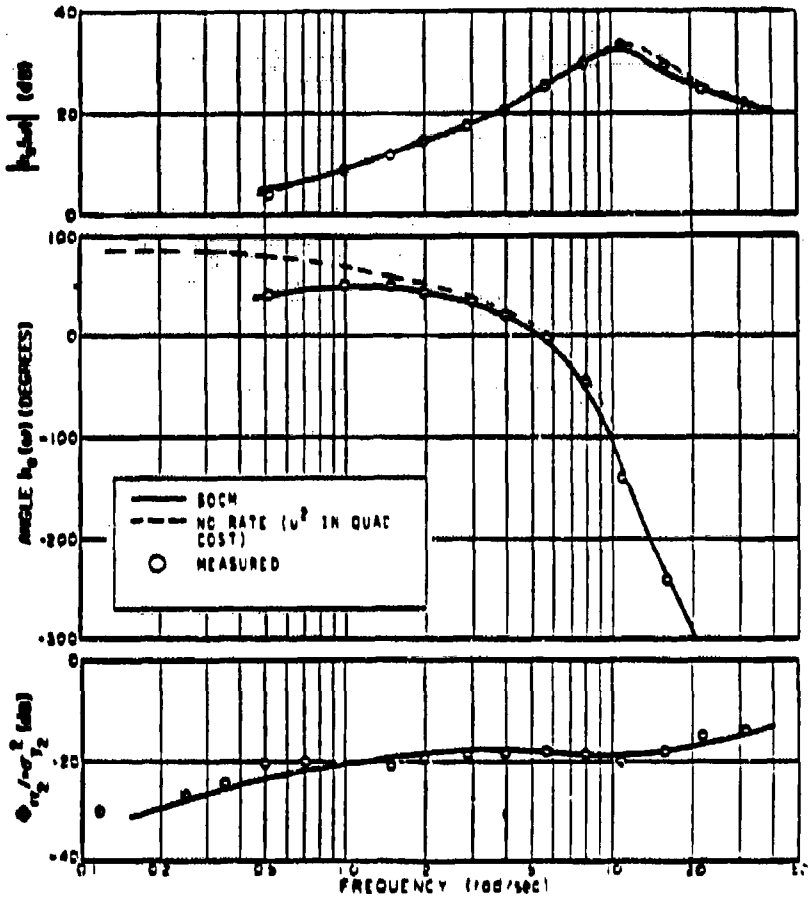


FIGURE 9. MEASURED AND PREDICTED FREQUENCY DOMAIN MEASURES k/s^2 DYNAMICS (Average of three subjects)

match to the k/s^2 data. The parameters corresponding to the fit are (No rate perception)

$$J = \int (x_2^2 + g\, u^2)\, dt$$

$g = 0.0005$

$T_N = 0.1057$

$\tau = 0.2$

$\rho_y = -20$ dB

$\rho_u = -25$ dB

TABLE III. ALTERNATE OCM STRUCTURES

NO RATE PERCEPTION

MODEL PARAMETER	k	k/s	k/s^2
CONTROL RATE WEIGHTING ρ	0.01	0.0005	0.0001
MOTOR TIME CONSTANT T_N	0.1112	0.1057	0.1077
TIME DELAY τ	0.15	0.17	0.20
MOTOR NOISE/SIGNAL RATIO ρ_m (dB)	-25	-20	-25
OBSERVATION NOISE/SIGNAL RATIO ρ_y (dB)	-50	-50	-40
MEAN SQUARE OUTPUT - MEASURED	0.13	0.13	0.014
MEAN SQUARE OUTPUT - MODEL	0.11	0.114	0.0104
MEAN SQUARE OUTPUT RATE - MEASURED	4.8	3.1	0.10
MEAN SQUARE OUTPUT RATE - MODEL	11.02	2.89	0.09
MEAN SQUARE CONTROL - MEASURED	0.53	4.2	1.43
MEAN SQUARE CONTROL - MODEL	0.527	3.61	0.883

mean squared output $\overline{x_2^2} = 0.0084$ (measured = 0.014)

mean squared output rate $\overline{\dot{x}_2^2} = 0.1024$ (measured = 0.1)

Note that this formulation has a higher value of the observation noise/signal ratio ρ_y as compared to that in Table III where u^2 is used in the cost. This is not consistent with the remnant data below $\omega = 1.0$ rad/s where the normalized remnant is greater than -20 dB. This can be remedied by introducing a pseudo-observation noise with the noise/signal ratio given above and a driving observation noise that is negligible (i.e.: noise/signal ratio = $-\infty$. The remnant fit requires the introduction of a motor noise filter as in the previous case.

VI. CONCLUSIONS AND RECOMMENDATIONS

The results in this paper have shown the existence of alternate human operator optimal control model structures capable of describing standard laboratory describing function and remnant data for k, k/s and k/s^2 plant dynamics. The alternate optimal control model formulation with no rate perception gives rise to almost negligible values for the observation noise/signal ratio ($\rho_y = -50$ dB). This implies that the Kalman filter component in the optimal control model can almost be treated as a Luenberger Observer or state reconstructor [21]. Further investigation into this possibility is underway and appears promising. Ultimately, the observer/optimal controller model may be further reduced to a lower order and eventually converge to the classical describing function/remnant human operator model.

In summary, the optimal control modeling approach is an attractive concept that provides a valuable tool for analyzing information processing and control aspects of human operator behavior. However, it is important to continually reexamine established tenets to avoid further embarassment. A questioning research philosophy where a given experiment is repreated by several investigators across continents is standard procedure in the natural sciences. This is possible because experimental data and methods are well documented for archival purposes. Consequently, investigators can compare results and resolve differences. Such an approach is reccommended in human operator research and technology, whereby data and methods would be systematically documented and stored in a library for future reference and reexamination. This can only help improve confidence in prevailing methodology and models, and remedy obvious drawbacks early.

REFERENCES

1. D. Kleinman, S. Baron and W.H. Levison; "An Optimal Control Model of Human Response - I: theory and validation". Automatica 6,357-370 (1970).

2. S. Baron, D.L. Kleinman and W.H. Levison; "An Optimal Control Model for Human Response - II: prediction of human performance in a complex task". Automatica 6,371-384 (1970).

3. D. Kleinman, S. Baron and W.H. Levison; "A Control Theoretic Approach to Manned-Vehicle Systems Analysis". IEEE Trans. Aut. Control AC-16, 824-832 (1971).

4. D.L. Kleinman and S. Baron; "Manned Vehicle Systems Analysis by Means of Modern Control Theory", Bolt Beranek and Newman, Inc., Report No. 1967, NASA ERC, Cambridge, Mass., Contract No. NAS12-104, June 1970.

5. S. Baron, et al., "Application of Optimal Control Theory to the Prediction of Human Performance in a Complex Task", Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio 45433, AFFDL-TR-69-81, March 1970.

6. W.H. Levison, J.I. Elkind and J.L. Ward; "Studies of Multivariable Manual Control Systems: A Model for Task Interference", Bolt Beranek and Newman, Inc., Report No. NASA CR-1746, May 1971.

7. G.E.P. Box and G.M. Jenkins, "Time Series Analysis", Holden-Day, San Francisco, 1976.

8. P. Whittle; "Tests of Fist in Time Series", Biometrica, 39, pp309-318.

9. R.L. Kashyap and A.R. Rao; "Dynamic Stochastic Models from Empirical Data" New York; Academic, 1976.

10. G.M. Jenkins and P.G. Watts; "Spectral Analysis and its Applications", Holden-Day, San Francisco, 1968.

11. W.H. Levison, S. Baron and A.M. Junker; "Modeling the Effects of Environmental Factors on Human Control and Information Processing", AMRL-TR-76-74, August 1976.

12. W.H. Levison and A.M. Junker; "A Model for the Pilots use of Motion Cues in Roll-Axis Tracking Tasks", AMRL-TR-77-40, June 1977.

13. K.M. Doyle and W.C. Hoffman; "Pilot Modeling for Manned Simulation, Volume II, Program User's Manual, AFFDL-TR 76-124, December 1976.

14. D.T. McRuer and H.R. Jex; "A Review of Quasi-Linear Pilot Models", IEEE Trans. Huaman Factors Electronics, Vol. HFE-8, No. 3, pp 231-249, September 1967.

15. R.J. Wasicko, D.T. McRuer and R.E. Magdaleno; "Human Pilot Dynamic Res-

ponse in Single-Loop Systems with Compensatory and Pursuit Displays", AFFDL-TR-66-137, December 1966.

16. A.V. Phatak, H. Weinert, I. Segall and C.N. Day; "Identification of a Modified Optimal Control Model for the Human Operator", Automatica 12, pp 31-41, 1976.

17. D.L. Kleinman; "Comments on the Application of Modern Control Theory to Manual Control". Systems, Man and Cybernetics Review, February 1977.

18. A.V. Phatak; "Formulation and Validation of Optimal Control Theoretic Models for the Human Operator". Systems, Man and Cybernetics Review, February 1977.

19. S. Baron; "Some Comments on Parameter Identification in the Optimal Control Model". Systems, Man and Cybernetics Review, February 1977.

20. N.K. Gupta and R.K. Mehra; "Computational Aspects of Maximum Likelihood Estimation and Reduction in Sensitivity Function Calculations", IEEE Trans. Auto Control, Vol AC-19, No. 6, December 1974.

21. D.G. Luenberger;" An Introduction to Observers", IEEE Trans. in Auto Control, Vol AC-16, December 1971.

A STRUCTURAL MODEL OF THE ADAPTIVE HUMAN PILOT

Ronald A. Hess*
Ames Research Center, NASA, Moffett Field, California

Abstract

A compensatory tracking model of the human pilot is offered which attempts to provide a more realistic representation of the human's signal processing structure than that which is exhibited by pilot models currently in use. Two features of the model distinguish it from other representations of the human pilot. First, proprioceptive information from the control stick or manipulator constitutes one of the major feedback paths in the model, providing feedback of vehicle output rate due to control activity. Implicit in this feedback loop is a model of the vehicle dynamics which is valid in and beyond the region of crossover. Second, error-rate information is continuously derived and independently but intermittently controlled. An output injected remnant model is offered and qualitatively justified on the basis of providing a measure of the effect of inaccuracies such as time variations in the pilot's internal model of the controlled element dynamics. The data from experimental tracking tasks involving five different controlled element dynamics and one nonideal viewing condition were matched with model generated describing functions and remnant power spectral densities.

Introduction

Since the end of World War II, flight control system engineers have devoted considerable energy to obtaining analytical descriptions of the process by which the human pilot exerts control over an aircraft in closed-loop tracking tasks. The analytical techniques for pilot modeling which have evolved in the intervening years have tended to fall into two categories, a dichotomy shared with control system design techniques in general. These are the so-called "classical" and "modern" approaches. The classical approach has led to the development of the "servo model" of the human pilot.[1] In this model, a servolike operation on a single stimulus (system error) by the pilot is hypothesized. The modern approach has yielded the "optimal control model" of the human pilot[2] in which the pilot's dynamic characteristics are likened to those of an optimal state estimator and regulator, and the pilot is assumed to operate on a vector stimulus of error and error rate. Both of these models have their particular merits and practitioners and both have been applied to similar analysis/design problems with reasonable success. These applications range from describing pilot dynamics in single-axis tracking tasks,[1,2] to cockpit display design,[3,4] to handling qualities investigations,[5,6] and to motion cue research.[7,8]

Despite their relative successes, neither the servo model nor the optimal control model attempts to describe the underlying structure which contributes to human pilot dynamics. Rather, these models have evolved as expedient means for quantifying the general transfer and performance characteristics of the human pilot in specific tracking tasks. Following the lead of Smith,[9] Hess[10] has attempted to provide a more satisfactory structural model of the human pilot. The impetus behind the research of Refs. 9 and 10 and that reported here is the conviction that a pilot model which provides a more realistic representation of the signal processing structure of the human pilot will also provide a more unified theoretical framework within which to interpret a variety of empirical pilot/vehicle response phenomena. The particular phenomena to be discussed here will include the ability of the pilot to adapt to different vehicle dynamics and to adapt to displays of varying quality. As used here, "adaptation" refers to the ability of the pilot to change his dynamic characteristics, through training, to suit the task at hand.

The model which is the subject of this paper is an outgrowth of the research described in Refs. 9-11. A block diagram representation of the model, suitable for describing such single-axis pilot behavior as pitch attitude tracking in the presence of atmospheric turbulence, is shown in Fig. 1. While the model structure may appear more complex than other single-axis pilot representations such as the servo model, it will be shown that, given the vehicle dynamics, only five of the model parameters in Fig. 1 need to be varied for the model to provide a reasonable match to experimentally derived pilot dynamic characteristics for a variety of vehicle dynamics and display conditions.

There are two features of this model which distinguish it from other representations of the human pilot. First, proprioceptive information from the control stick or manipulator constitutes one of the two major feedback paths in the model, visual (and/or aural or tactile) error sensing constituting the other. This proprioceptive information is used to provide an estimate of the rate of change of the pertinent vehicle output quantity due to pilot control inputs, e.g., an estimate of aircraft pitch rate due to longitudinal stick inputs. Second, error-rate information is continuously derived from visual (aural or tactile) stimuli and is independently controlled. The control is not continuous, however (note the switch in Fig. 1). As opposed to other models which utilize continuous error and error-rate information to create a desirable open-loop transfer function[1,2] (i.e., a combined pilot/vehicle transfer function of K/s in the region of crossover), the model discussed here assumes that error rate is controlled independently. This implies that the pilot is essentially a single-channel processor in terms of information derived from visual (aural or tactile) sensors and that he switches between error and error-rate control during the tracking task. As will be shown, error control dominates.

Model Specifics

The model of Fig. 1 has been divided into "central nervous system" and "neuromuscular system" components, a division which is intended to emphasize the nature of the signal processing activity involved.

---

*Research Scientist. Member AIAA.

This paper is declared a work of the U.S. Government and therefore is in the public domain.

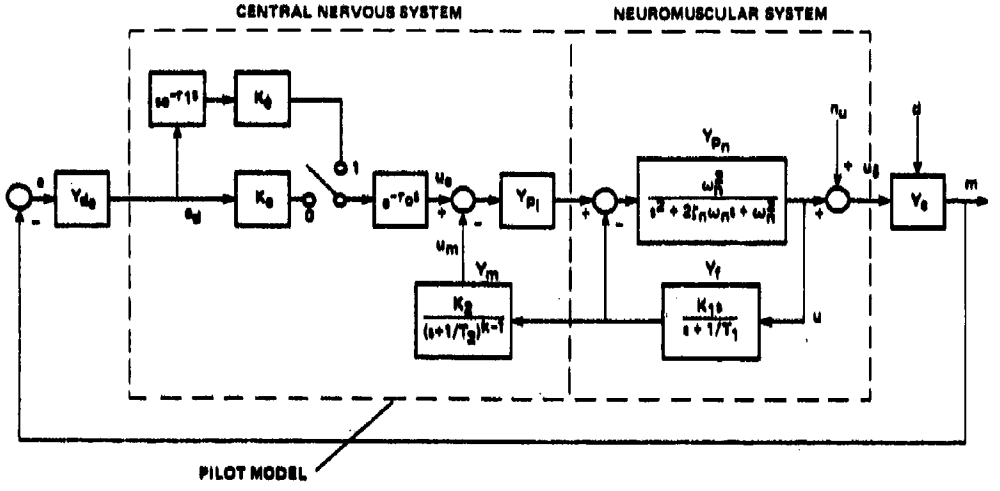


Fig. 1 Structural model of the adaptive human pilot.

System error $e(t)$ is presented to the pilot via a display with dynamics Y_{d_p}. The rate of change of the displayed error is assumed to be derived from $e_d(t)$. The process of deriving error rate is assumed to entail a computational time delay of τ_1 seconds. Constant gains K_0 and K_1 multiply the signals $e_d(t)$ and $\dot{e}_d(t - \tau_1)$, respectively. The switch allows either of these two signals to be used as driving signals to the remainder of the model. In this study, the action of the switch will be parameterized by the variable P_1, which represents the probability that the switch will be in position 1 (error-rate control) at any instant of time. A central time delay of τ_c seconds is included to account for the effects of latencies in the visual (aural or tactile) process sensing $e_d(t)$, motor nerve conduction times, etc. The resulting signal $u_c(t)$ provides a command to a closed-loop system consisting of "pulsing logic" Y_{P_1}, a model of the open-loop neuromuscular dynamics of the particular limb driving the manipulator, Y_{P_n}, and elements Y_f and Y_m which emulate, at least approximately, the combined effects of the muscle spindles and the dynamics associated with higher level signal processing. A colored noise $n_u(t)$ is injected at the pilot's output as remnant.

Let us look at each of the model elements now in a bit more detail. The assumption that error rate is a derived rather than a sensed quantity is conservative in nature. In terms of visual stimuli, arguments still abound as to whether motion in the visual field of the human is perceived "directly" as a primary psychological attribute of the visual system, or whether it is inferred or is a calculated or secondary attribute based on memory for position.[12] In terms of the model, the question of direct vs derived rate inputs will influence the magnitude of the time delay τ_1; i.e., for direct rate perception τ_1 will be negligible, whereas for derived rate inputs a significant τ_1 may be indicated. Since, here, the stimulus $e_d(t)$ can be visual, aural or tactile, it was assumed that $\dot{e}_d(t)$ was a derived quantity. A simple model for such rate derivation was offered by McRuer et al.[13] in the form of a "differential-displacement" model. As a conservative estimate of the sampling delay occurring in such a model, let us use $\tau_1 = 0.2$ sec. The time delay τ_c can be broken down into constituent components, at least in approximate fashion. For visual stimuli, which will constitute the modality for the data to be utilized later, we have the following average values:[13]

Latency of the visual process, sec	0.075
Motor nerve conduction time, sec	0.030
Central processing time, sec	0.030
	0.135

The value $\tau_c = 0.14$ sec was used here.

A discussion of appropriate pulsing logic was carried out previously by Hess[11] and will not be dealt with here. This logic is used to generate the pulsive control behavior evident when the pilot is faced with higher order controlled element dynamics such as K/s^2 or K/s^3. In this study, $Y_{P_1} = 1.0$. The form of the open-loop neuromuscular dynamics, Y_{P_n}, is based upon electromyogram measurements reported in Ref. 14. The following values of ω_n and ζ_n were selected based upon that work: $\omega_n = 10$ rad/sec, $\zeta_n = 0.707$.

The feedback loop in Fig. 1 involving Y_f alone is intended to represent, in approximate fashion, the feedback activity of the muscles spindles in the limb driving the manipulator or control stick. For the purposes of this study, $K_1 = 1.0$. The washout characteristics of the element Y_f are very important to the overall model dynamic characteristics, as will be seen. The physiological source of such a washout, if it indeed exists, is unknown. Its existence is part of the basic structural hypothesis

of the model. From a functional standpoint, the muscle spindle is a length transducer transforming increases in extra-fusal muscle fiber length from some set or operating point into electrical impulses. Anatomically, the spindles are in parallel with the extra-fusal fibers. One possibility is that the washout characteristics are a consequence of the intrafusal fibers of the spindle tracking the low frequency changes in the length of extra-fusal fibers in the agonist-antagonist muscle pairs in the limb driving the manipulator. A necessary condition for such coactivation of muscle fibers would be concomitant coactivation of the motor fibers which innervate the extra- and intra-fusal fibers, respectively. Such a coactivation has been documented by physiologists and given the name "alpha-gamma linkage."[15]

The form of the element Y_p is assumed to depend on the order of a simplified model of the controlled-element dynamics valid in the frequency range where $Y_p \cdot Y_c \triangleq 1.0$, i.e., open-loop crossover. For all but two of the controlled elements to be studied here, no simplification is necessary. In general, $\hat{Y}_c(s) \triangleq Ks^k$, $k = 0, \pm 1, \pm 2, \ldots$, etc., where $\hat{Y}_c$ is an approximation to Y_c valid in the region of open-loop crossover.[10,11] Note that manipulator dynamic characteristics can be lumped into Y_c. Evidence supporting the hypothesis of a proprioceptive-related internal model has recently been offered by Kessel and Wickens.[16] The use of the factor $(s + 1/T_2)^k$ in Y_m ($T_2 \triangleq T_1$) rather than simply s^k was found to be necessary to achieve acceptable describing function matches with experimental data. The inclusion of this term is tantamount to saying that no low frequency processing of spindle output takes place in the higher levels of the nervous system, only spindle output itself is utilized, attenuated or magnified by an appropriate gain K_2. For the sake of simplicity, $T_2 = T_1$ in all that follows. Note that in the region of open-loop crossover, the signal $u_m(t)$ is an estimate of the rate of change of the controlled-element output due to control activity $u_c(t)$, excluding the contributions due to disturbance $d(t)$ and remnant $n_u(t)$. Thus, this loop is a form of rate feedback in the model.

The physical interpretation of remnant depends on the particular pilot model with which it is associated. The servo model had its origins in nonlinear describing function theory, and remnant for that model tends to be thought of, indeed defined, as that portion of the pilot's output not linearly correlated with the disturbance. The tacit interpretation here is that remnant arises from nonlinearities and/or time variations in the human pilot. In the optimal control model, which obviously had its origins in linear optimal estimation and control theory, remnant tends to be viewed in terms of "observation noise." Here the interpretation is that all observations which the pilot makes are corrupted by white noise. This is not to say that there is no underlying equivalence in these views[17] but rather that the physical interpretation of remnant is often determined by the genesis of the pilot model itself. This is as it should be. In the model of Fig. 1, the pilot's internal representation of the controlled-element dynamics is hypothesized to form a pivotal part of the equalization capability of the pilot. Thus, inaccuracies, such as time variations, in this internal model representation can be logically considered a primary source of remnant. Furthermore, this remnant should be amenable to representation as a "process noise" in that it is intended to account, in stochastic fashion, for the effects of imprecise internal model parameterization,[18] whatever the source of this imprecision. Thus, in Fig. 1, the remnant $n_u(t)$ is shown injected into the pilot's output and, as such, acts as a true process noise in terms of the vehicle to be controlled.

With the above discussion in mind, two hypotheses will be made regarding the nature of this injected noise. The first hypothesis asserts that the variance of $n_u(t)$ scales with the variance of the system error normalized by the manipulator/controlled-element static gain. This means the variance of the noise scales with $\sigma_{\delta_d}^2/K^2|Y_{d_e}(j\omega)|^2$, where $\sigma_{\delta_d}^2$ is the variance of the displayed error, $Y_{d_e}(j\omega)$ is the display dynamics (typically a gain) and K is the manipulator/controlled-element static gain. The second hypothesis asserts that the process noise power is attenuated at frequencies beyond the bandwidth of the human pilot. This limitation can most easily be accommodated in the model by considering $n_u(t)$ to be a colored noise with a bandwidth determined by the position of the closed-loop neuromuscular mode roots of the describing function portion of the pilot model.

The idea that the variance of $n_u(t)$ should scale with normalized system error means that $n_u(t)$ is not multiplicative noise in that it does not scale with the variance of the signal to which it is added. The first hypothesis is based on the reasonable supposition that gain-normalized system-error variance is a sensitive indicator of errors in the pilot's internal representation of the controlled-element dynamics. The assertion that the control-injected remnant variance scales linearly with this error variance is an assumption to be verified on the basis of experiment. The second hypothesis seems almost self-evident; there can be no significant effects attributable to internal model errors at frequencies well beyond the bandwidth of the structure which contains the internal model, i.e., the pilot.

The model for the power spectral density of the injected remnant $n_u(t)$ can now be given as

$$\Phi_{nn_u}(\omega) = \rho \cdot \pi \cdot \frac{\sigma_{\delta_d}^2}{K^2|Y_{d_e}(j\omega)|^2} \frac{(1/T_3)^2}{[\omega^2 + (1/T_3)^2]}$$

where ρ is a scale factor to be empirically determined and $1/T_3$ represents the undamped natural frequency of the closed-loop neuromuscular mode roots of the describing function portion of the pilot model. On the basis of remnant data for pure gain, controlled-element dynamics, a value of $\rho = 0.119$ was selected via the data fitting procedure to be described. This value was then used in the remnant model for <u>all</u> the controlled-element dynamics studied herein.

By referring to Fig. 1, one can see that, with the switch in position 1, the time delay for error-rate tracking is $\tau_0 + \tau_1 = 0.34$ sec. The magnitude of this delay will obviously result in a compromise between responsiveness and stability for the rate tracking loop. Since error rather than error-rate tracking will predominate in the model, it was felt that the responsiveness of the error-rate loop should take precedence over relative stability characteristics. Therefore, in the data fitting procedure to

be described, the gain K_g will be made as large as possible, under the single constraint of guaranteeing only absolute stability in the rate tracking loop.

Data Fitting

As employed here, the model of Fig. 1 will be a "data summary" rather than a "predictive" model of the human pilot. We will not attempt to make predictions about the behavior of the pilot/vehicle system under situations which have not been previously tested. Here "data" is intended to mean the frequency domain measures of describing function and remnant power spectral density. In the available experimental data to be utilized in the fitting or matching procedure, the pilot model remnant was assumed to be injected at the pilot's input rather than his output. Therefore, in order to compare the experimental and model remnant spectra, we will use the output injected remnant model just discussed to obtain equivalent input (error) injected spectra. Let us now turn our attention to such a data matching procedure.

The following parameters will be at our disposal in varying the frequency domain characteristics of the structure of Fig. 1: the gains K_0, K_1, and K_2, the time constant T_1, and the error-rate control probability P_1. The remaining variables: k, K_1, T_2, T_3, ζ_n, ω_n, τ_0, τ_1 and ρ will either be obvious from the task (i.e., k is determined once Y_c is specified) or be functions of the former five variables, e.g., $T_2 = T_1$, $T_3 = f(\zeta_n, \omega_n, K_2, T_1)$ or simply be held fixed at empirically determined values. The decision of which variables to vary and which to hold fixed is not entirely arbitrary. The variables τ_0, τ_1, ζ_n, ω_n, K_1 and to some extent ρ are intended to be strongly isomorphic with the physiology of the human pilot, at least at the level of detail possible in the structure of Fig. 1. While it is not difficult to imagine rather small intertask variations in these parameters, large variations are somewhat less palatable, at least to this author. Said another way, one should not be forced to consistently hypothesize 25 to 50% variations in basic (as opposed to "effective") time delays or open-loop neuromuscular parameters in order to interpret pilot response phenomena. On the other hand, the variables K_0, K_g, K_2, P_1 and to some extent T_1 represent the "central processing" part of the structure of Fig. 1. Large variations in these parameters from task to task are certainly not unreasonable.

Were it not for the existence of the switch in Fig. 1, calculating model describing functions and remnant spectra would be a simple exercise in block diagram algebra. The switch and its operation form an important part of the model, however, and represent a significant departure from the model of Refs. 10 and 11. Given the model with some fixed set of parameters, the question to be answered is: what describing function $Y_p(j\omega)$ and error-injected remnant power spectral density $\Phi_{nn_{e_d}}(\omega)$ would be obtained by making spectral measurements of $u_g(t)$ and $e_d(t)$ over some fixed run length with the switch randomly moving between positions 0 and 1 and with the probability of being at position 1 of P_1? To answer this question, one could, of course, simulate the model and actually make the measurements using typical spectral techniques such as the Fast Fourier Transform. However, this would be computationally quite expensive, particularly since many such runs would be required to select the proper model parameters to fit the data. Consider, instead, the following approximate development. Let $u_g(t)$ and $e_d(t)$ represent the time histories of model input and output in a single run with the switch operating as just described. Concentrating on $u_g(t)$, let $Fu_g(j\omega,P_1)$ represent a spectral measure (Fourier coefficient or power spectral density) of $u_g(t)$. Expanding $Fu_g(j\omega,P_1)$ in a Taylor series about the point $P_1 = 0$:

$$Fu_g(j\omega,P_1) = Fu_g(j\omega,0) + \left.\frac{\partial Fu_g(j\omega,P_1)}{\partial P_1}\right|_{P_1=0} P_1 + \left.\frac{\partial^2 Fu_g(j\omega,P_1)}{\partial P_1^2}\right|_{P_1=0} P_1^2 + \text{H.O.T.}$$

Now since $P_1 < 1$ (for the cases to be studied here $P_1 \leq 0.25$) we can write

$$Fu_g(j\omega,P_1) \triangleq Fu_g(j\omega,0) + \left.\frac{\partial Fu_g(j\omega,P_1)}{\partial P_1}\right|_{P_1=0} P_1$$

Further, let us approximate

$$\left.\frac{\partial Fu_g(j\omega,P_1)}{\partial P_1}\right|_{P_1=0} \triangleq \frac{Fu_g(j\omega,1) - Fu_g(j\omega,0)}{(1-0)}$$

Therefore, we can state

$$Fu_g(j\omega,P_1) \triangleq (1-P_1)Fu_g(j\omega,0) + P_1 Fu_g(j\omega,1)$$

For efficiency, let us denote the right-hand side of the equation above as $[Fu_g(j\omega)]'$, dropping the P_1 notation for convenience. Thus, we have

$$Fu_g(j\omega) \triangleq \left[Fu_g(j\omega)\right]' = (1-P_1)Fu_g(j\omega,0) + P_1 Fu_g(j\omega,1)$$

A similar relation can, of course, be obtained for $Fe_d(j\omega)$.

The Appendix offers a simple but quantitative example of the quality of this approximation procedure. The Appendix also indicates that, in terms of frequency domain measures, the switching operation can be successfully parameterized by the single quantity P_1. Measured Fourier coefficients appear to be relatively insensitive to the minimum duration of switch closure for a range of values (0.3 to 5.0 sec) deemed appropriate for the tracking tasks to be utilized here.

Now with a sum of sinusoids disturbance, the model describing function $Y_p(j\omega_k)$ which would be measured in the single run with the switch operating randomly can be approximated as:

$$Y_p(j\omega_k) = \frac{C_{u_g}(j\omega_k)}{C_{e_d}(j\omega_k)} \triangleq \frac{[C_{u_g}(j\omega_k)]'}{[C_{e_d}(j\omega_k)]'} \qquad (1)$$

With remnant $n_{e_d}(t)$ injected at the displayed error $e_d(t)$, we have, for the power spectral density of the model output at frequencies ω_j:

$$\Phi_{uu_\delta}(\omega_j) = |C_{u_\delta}(j\omega_j)|^2 = \frac{\Phi_{nn_{e_\delta}}(\omega_j)\left|\frac{C_{u_\delta}(j\omega_j)}{C_{e_\delta}(j\omega_j)}\right|^2}{\left|1 + \frac{C_{u_\delta}(j\omega_j)}{C_{e_\delta}(j\omega_j)} \cdot Y_c(j\omega_j)\right|^2} \quad (2)$$

where the ω_j occur between disturbance frequencies ω_k and $C_{u_\delta}(j\omega_j)$ and $C_{e_\delta}(j\omega_j)$ are obtained by averaging their values at the disturbance frequencies to either side of ω_j. With remnant $n_u(t)$ injected at the model output we can write

$$[\Phi_{uu_\delta}(\omega_j)]' = [|C_{u_\delta}(j\omega_j)|^2]' = \left[\frac{\Phi_{nn_u}(\omega_j)}{\left|1 + \frac{C_{u_\delta}(j\omega_j)}{C_{e_\delta}(j\omega_j)} \cdot Y_c(j\omega_j)\right|^2}\right]' \quad (3)$$

Again, the measured power spectral density of the model output $u_\delta(t)$ with the switch in operation can be approximated as

$$\Phi_{uu_\delta}(\omega_j) \triangleq [\Phi_{uu_\delta}(\omega_j)]' \quad (4)$$

Thus:

$$\Phi_{nn_{e_\delta}}(\omega_j) \triangleq \left[\frac{\Phi_{nn_u}(\omega_j)}{\left|1 + \frac{C_{u_\delta}(j\omega_j)}{C_{e_\delta}(j\omega_j)} \cdot Y_c(j\omega_j)\right|^2}\right]' \cdot \frac{|1 + Y_p(j\omega_j)Y_c(j\omega_j)|^2}{|Y_p(j\omega_j)|^2} \quad (5)$$

Equations (1) and (3) will allow us to approximate the model describing functions and error-injected remnant power spectral densities which would be "measured" in a single run with the switch in operation by spectral measures obtained with the switch either in the 0 or 1 position for the entire run.

Finally, in order to avoid the necessity of actually making simulation runs to obtain the spectral measures, the complex number on the right-hand side of Eq. (1) will be expressed in magnitude and phase form as:

$$|Y_p(j\omega_k)| = \left|\frac{\left[C_{u_\delta}(j\omega_k)\right]'}{\left[C_{e_\delta}(j\omega_k)\right]'}\right| \triangleq \log^{-1}\left\{\left[\log\left|\frac{C_{u_\delta}(j\omega_k)}{C_{e_\delta}(j\omega_k)}\right|\right]'\right\} \quad (6)$$

$$\angle Y_p(j\omega_k) = \angle \frac{\left[C_{u_\delta}(j\omega_k)\right]'}{\left[C_{e_\delta}(j\omega_k)\right]'} \triangleq \left[\angle \frac{C_{u_\delta}(j\omega_k)}{C_{e_\delta}(j\omega_k)}\right]' \quad (7)$$

where $C_{u_\delta}(j\omega_k)/C_{e_\delta}(j\omega_k)$ is now merely the algebraically obtained expression for the model transfer function with the switch in either the 0 or 1 position. Equations (6) and (7) thus allow simple algebraic averages in the frequency domain to replace Fourier transformations in the time domain. The accuracy of the relations on the right-hand side of Eqs. (6) and (7) was appraised by creating two sets of complex numbers A and B, whose real and imaginary parts were randomly selected from a population with a Gaussian distribution possessing a mean of 0.3 and a standard deviation of unity. Each set contained 5,000 complex numbers. The "primed" operations on the middle and right-hand sides of Eqs. (6) and (7) were generalized by considering the expected values of A and B (denoted by E[A], E[B], respectively). Values for $|E[A]/E[B]|$ and $\angle E[A]/E[B]$ corresponding to an arbitrarily large sample size would be 1.0 and 0°, respectively. The actual values generated for a sample size of 5,000 were:

$$\left|\frac{E[A]}{E[B]}\right| = 1.009, \qquad \log^{-1}\left\{E\left[\log\left|\frac{A}{B}\right|\right]\right\} = 1.01$$

$$\angle \frac{E[A]}{E[B]} = -1°, \qquad E\left[\angle \frac{A}{B}\right] = 1.75°$$

The expressions on the right-hand sides of Eqs. (6) and (7) are seen to provide very satisfactory approximations for our purposes.

The validity of the approximate analysis just described obviously depends upon the validity of Eqs. (1) and (4) and the relations which preceded them. While these relations are intuitively appealing, at least to the author, they have not been rigorously established. The problem, of course, stems from the fact that the model proposed here is fundamentally nonstationary in nature due to the presence of the switch. The archival data dealing with human pilot dynamics, on the other hand, have been obtained using spectral measurement techniques which tacitly assume that the human is fundamentally stationary in his tracking behavior. Thus, one is presented with something of a dilemma in attempting to select parameters in a model like that of Fig. 1 to match archival data. Suffice to say, the analysis just described offers an approximate means of doing this. As will be seen, the nonstationary aspects of the model are mitigated by the fact that error rather than error-rate control dominates, i.e., the parameter selection procedure to be described indicates $P_1 \leq 0.23$ for all the tasks studied.

After the remnant scale factor $\rho = 0.119$ had been obtained on the basis of remnant data for pure gain dynamics, data fits for the remaining controlled-element dynamics and nonideal viewing condition were obtained in a straightforward manner. First a particular set of experimental data was selected. This allowed Y_c, k, $Y_{e_\delta}(j\omega)$ and K to be specified immediately. Next, with the parameters K_1, τ_0, τ_1, ζ_n, ω_n and ρ fixed at the values shown in the first row of Table 1, P_1 was set to zero and K_2 was set to unity. Next, T_2 and K_2 were varied until the model describing function provided a close match to the shape of the experimental describing function at and beyond the crossover frequency. Next, the value of K_1 corresponding to neutral closed-loop stability for error-rate tracking (with the switch in the 1 position) was found. With this value, P_1 was incremented from zero until the shape of the low frequency model describing function matched the data. The gain K_c was then adjusted so that the 0 dB line of the model describing function amplitude ratio matched the data. In all of these cases, the quality of the fit was determined by eye, i.e., no formal numerical criterion was employed.

Once K_2 and T_2 were selected, the closed-loop neuromuscular mode roots were available and the control-injected remnant model was complete. The error-injected remnant was calculated using Eq. (5) after an acceptable describing function fit was obtained. Thus, no iteration on the remnant model was necessary. The error-injected remnant power spectral density was normalized by the experimental displayed-error variance. Of course, one could have

Table 1 Pilot model parameter values used to generate describing functions and remnant power spectral densities in Figs. 2-7

Controlled element dynamics	Model parameters											
	k	K_e	K_d	K_2	P_1	T_1	K_1	τ_0	τ_1	ζ_n	ω_n	ρ
K	0	11.1	2.13	2.0	0.05	5.0	1.0	0.14	0.2	0.707	10.0	0.119
K/s	1	22.2	3.42	2.0	0.05	5.0	1.0	0.14	0.2	0.707	10.0	0.119
K/s^2	2	26.2	10.48	10.0	0.20	2.5	1.0	0.14	0.2	0.707	10.0	0.119
K/s(s - 1)	2	89.6	28.6	30.0	0.20	1.0	1.0	0.14	0.2	0.707	10.0	0.119
$K/s(s^2 + 1.41s + 1)$	3	115.5	13.0	35.0	0.20	0.85	1.0	0.14	0.2	0.707	10.0	0.119
$K/s\|_{peripheral}$	1	12.6	2.52	0.75	0.25	5.0	1.0	0.14	0.2	0.707	10.0	0.119

used the model to generate $\sigma^2_{e_d}$. This would have been done after the parameters K_2, T_2, K_e, K_d, and P_1 had been selected and in an iterative manner similar to that employed in the optimal control model of the human pilot.[2] The experimental variance was used for two reasons. First, it obviated the iterative computation just described. Second, and more important, error variance is a redundant measure of model accuracy once a reasonable fit to experimental describing function and remnant is obtained.

Figures 2 through 7 show the model generated and experimental describing functions and remnant spectral densities for controlled elements with dynamics of $Y_c = K$, K/s, K/s^2, $K/s(s - 1)$, $K/s(s^2 + 1.414s + 1)$ and K/s with 22° peripheral viewing. Data for the first three dynamics were taken from Ref. 2, the fourth and sixth dynamics from Ref. 19, and the fifth dynamics from Ref. 20. In all cases, the pilot's display was a CRT screen and system error was presented by the displacement of a moving line from a fixed reference position. Control stick dynamics were negligible. In those cases where the data assumed remnant injected into error-rate, an equivalent error-injected spectrum was calculated. Note that for remnant representation, the decibel is defined as $10 \cdot \log_{10}(\cdot)$. Table 1 summarizes the model parameters used to obtain the figures using the approximation scheme just outlined. The dynamics of the first four controlled elements are identical to those used in Ref. 10 where an earlier form of the pilot model was used to match the data. In each case, the fits obtained here are superior to those of Ref. 10, and fewer model parameters need to be varied in order to obtain these fits (eight parameters in Ref. 10 as opposed to five here).

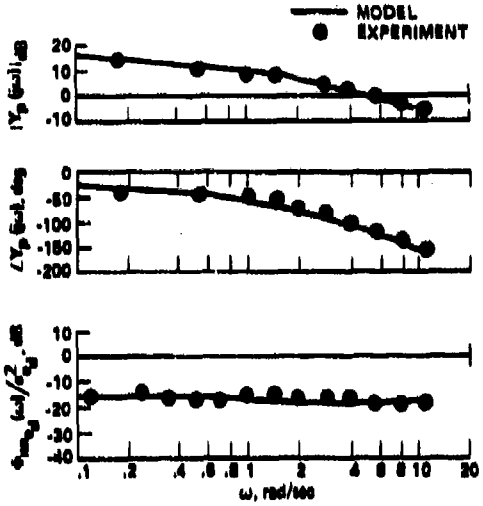


Fig. 2 Model and experimental describing functions and injected remnant power spectral densities, K controlled element dynamics.

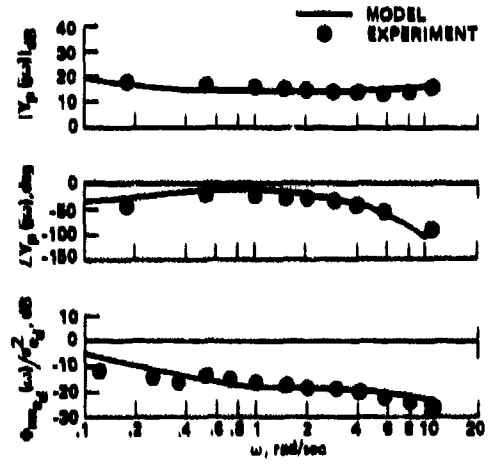


Fig. 3 Model and experimental describing functions and injected remnant power spectral densities, K/s controlled element dynamics.

Discussion

Adapting to Vehicle Dynamics

$Y_c = K$ (Fig. 2). The model and experimental describing function exhibits the first-order lag characteristics one would expect when pure gain dynamics are being controlled. Of particular interest is the model's ability to match the low frequency phase lag apparent in the data. The earlier form of the model discussed in Ref. 10 did not provide a satisfactory match to this low frequency phase data. The model and experimental injected remnant spectra are quite flat in the frequency range of interest. As mentioned previously, this experimental remnant was used to obtain the ρ value shown in Table 1.

$Y_c = K/s$ (Fig. 3). Here the model and experimental describing functions exhibit the pure gain

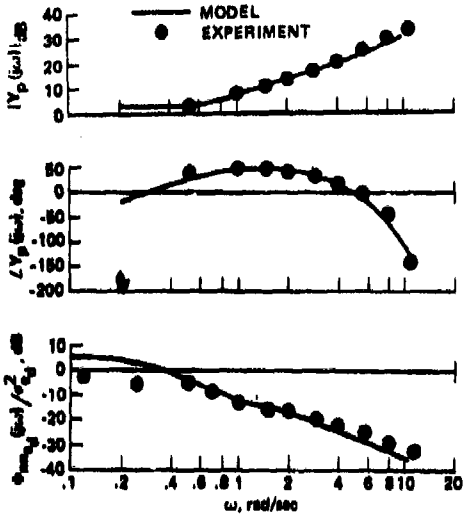


Fig. 4 Model and experimental describing functions and injected remnant power spectral densities, K/s^2 controlled element dynamics.

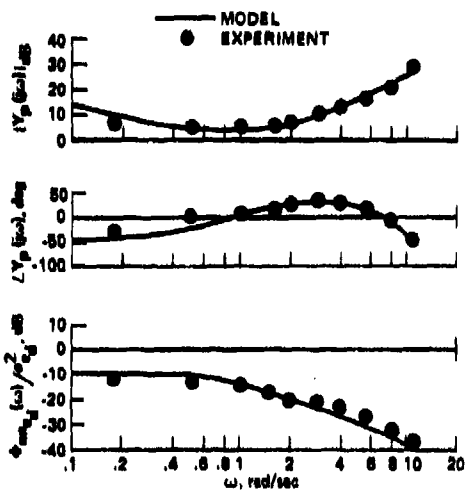


Fig. 5 Model and experimental describing functions and injected remnant power spectral densities, $K/s(s - 1)$ controlled element dynamics.

characteristics associated with rate dynamics. Again note the model's ability to match the low frequency phase lag (or phase "droop") in the data. The remnant match is particularly interesting in that the model is indicating a low-frequency rise in remnant power. In fact, rather than resembling the first-order process suggested by Levison et al.,[17] the remnant looks more like that offered by Pew:[21]

$$\Phi_{nn_{ed}}(\omega)/\sigma^2_{ed} = 0.04 + (\omega)^{-.8}$$

The ability of measured remnant spectra to exhibit shapes not amenable to description by a first-order

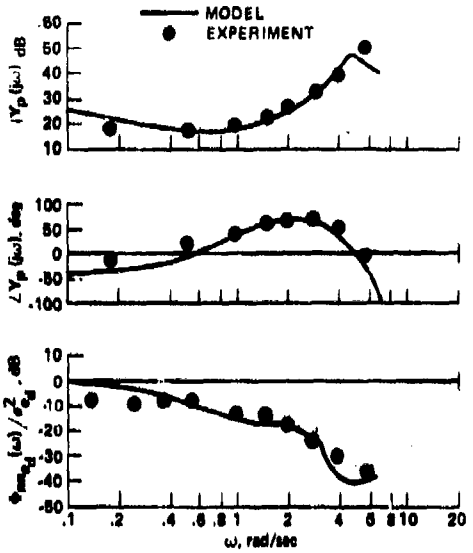


Fig. 6 Model and experimental describing functions and injected remnant power spectral densities, $K/s(s^2 + 1.414s + 1)$ controlled element dynamics.

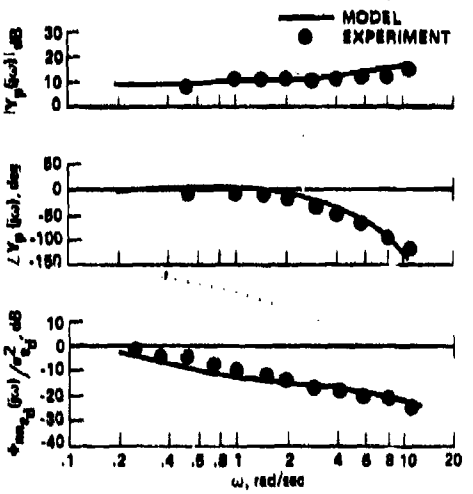


Fig. 7 Model and experimental describing functions and injected remnant power spectral densities, K/s controlled element dynamics with 22° peripheral viewing.

process has been noted by other researchers in addition to Pew, e.g., Jex, et al.[22] However, to the author's knowledge, this is the first time a model-generated remnant has exhibited these characteristics.

$Y_c = K/s^2$ (Fig. 4). Here, a first-order lead is evident in the describing function data. Calculation of the model transfer function with the switch in position 0 demonstrated that the lead is attributable to the closed-loop characteristics of the inner feedback loops in the model, not to the

error-rate utilization. The model-generated remnant spectrum now appears more like a first-order process. Here, for the first time, significant switching activity is involved in the model, i.e., $P_1 = 0.2$ as indicated in Table 1.

$Y_c = K/s(s - 1)$ (Fig. 5). The unstable dynamics again precipitate apparent lead equalization in the model. Again, however, this is due to inner loop activity, not error-rate utilization. In Ref. 10, the model structure had to be altered to allow for actual, continuous error-lead equalization to match this data.

$Y_c = K/s(s^2 + 1.414s + 1)$ (Fig. 6). These dynamics are interesting in that they resemble a K/s^3 plant in and beyond the region of crossover. Second-order apparent lead equalization is evident in model and data. The model exhibits more phase droop than the data, but, in general, the describing function fit is quite acceptable, particularly since Ref. 20 indicates that a substantial fraction of the pilot's output power was remnant related. It should be noted that the effective K/s^3 dynamics will probably induce pulsive or even impulsive control activity on the part of the pilot. Although no mention of such activity was made in Ref. 20, Jex and Allen[21] found that extremely pulsive control action accompanied "double-lead" equalization in a series of compensatory tracking tasks which they were studying. As mentioned previously, pulsive activity can be included in this model and is discussed in detail in Ref. 11. In terms of its effect on model describing function and remnant, the pulsive activity associated with this task will quite likely lower the apparent damping ratio of the closed loop neuromuscular mode roots. This would, in turn, improve the high frequency remnant and describing function matches of Fig. 6.

Comments

As Table 1 indicates, the generation of the remnant power spectral densities and the various lags, leads, low frequency phase droops and changes in high frequency phase lags evident in the overall model describing function u_s/e_d were obtained without changing the fundamental structure of the model or the values of the basic time delays (τ_0, τ_1), open-loop neuromuscular system parameters (ζ_n, ω_n and K_1), and model output injected remnant scale factor ρ. The pilot's adaptation to different controlled-element dynamics was accounted for by changes in the "internal model" of these dynamics, as reflected in changes in k, and changes in the parameters K_θ, $K_{\dot\theta}$, K_2, P_1 and T_1 primarily associated with information processing in that portion of the model representing higher levels of the central nervous system.

Adapting to Displays of Varying Quality

The experimental describing function and remnant of Fig. 7 correspond to experimental conditions identical to those which yielded the data of Fig. 2 save that the display was located 22° off the pilot's foveal axis at all times. Comparing the describing functions (both model and experimental) for the peripheral as opposed to the foveal viewing, one finds an increase in high frequency phase lags (or "effective" time delay), a decrease in low frequency phase lags, a decrease in crossover frequency and a concomitant increase in phase margin. Comparing the remnant (both model and experimental) for the two tasks reveals a significant increase in low frequency remnant power for the peripheral viewing case. Comparing the entries in the second and last rows of Table 1 indicates that these changes in the model-generated describing function and remnant in going from foveal to peripheral are attributable primarily to reductions in the gains K_θ and $K_{\dot\theta}$ and an increase in the probability of the switch being in position 1, P_1. An analysis of each of these factors shows that the increased error-rate tracking (P_1) is responsible for the increase in low frequency remnant power and decrease in low frequency phase lag. The decrease in $K_{\dot\theta}$ accounts for 70% of the increase in high frequency phase lag (the rest being attributable to the increase in P_1). Finally, the decrease in crossover frequency and increase in phase margin is due to the decrease in the gain K_θ.

A better fit to the low frequency remnant data could have been obtained by allowing $|Y_{d_\theta}(j\omega)|^2$ to attenuate as $\omega \to 0$. Since this term appears in the denominator of the expression for the variance of $n_u(t)$, an increase in control- or error-injected remnant power would occur. Attenuation of the display gain for low frequencies provides an analytical means for accounting for the experimentally verified fact that displays which are viewed peripherally offer poor low frequency error information.[20] For the sake of simplicity, this was not done here, however.

Fitting experimental data for this peripheral viewing experiment would be of little more than academic interest were it not for the fact that the changes in describing function and remnant just described have been shown to accompany other experimental studies involving degradations in display quality such as displays with "nonideal" quantized formats, etc.[22] Tactile displays can often be placed in this category. Recently, Schmid and Bekey[24] conducted tracking experiments involving an electrotactile display. One of their primary findings was that the particular electrotactile display which they utilized induced considerably larger high frequency phase lags (or effective time delays) in the describing functions of the test subjects as compared to those normally found in visual tracking with ideal display formats. Thus, the model parameter variations involved in matching data for the pair of tracking tasks involving foveal and peripheral viewing conditions may be useful in describing the manner in which the human adapts to any display degradation.

Speculation and Conclusions

This paper would not be complete without some speculation regarding the error-rate tracking characteristics of the model. As pointed out in the Introduction, the general frequency domain characteristics of the human pilot have been found to be very similar to the characteristics which a properly designed compensator would exhibit if it were to perform the same task as the pilot, regardless of whether the compensator characteristics were developed via classical or modern techniques. In this context, the role of the switch of Fig. 1 is unclear. As employed here, error-rate tracking is not necessary to create open loop transfer characteristics which a designer would deem desirable. This is accomplished via the hypothesized proprioceptive

feedback loops of Fig. 1. Indeed, error-rate tracking does not have a significant effect upon pilot model dynamics in and beyond the region of crossover.

Thus, faced with designing a compensator to maintain a controlled element at some equilibrium or trim point in the presence of disturbances of known spectral characteristics, error-rate tracking, per se, would probably not enter a designer's mind. The human pilot, however, views the control problem from a different vantage point than the designer. First, he operates in the time domain rather than the frequency domain. As a compensator, his dynamic characteristics are sometimes subject to significant intra-run variations. Lapses in attention and occasional control reversals can have a dramatic effect upon short term tracking performance and stability, especially with higher order controlled element dynamics. In addition, the pilot does not have the luxury of knowing the precise spectral characteristics of the disturbance. All of these factors make the independent, albeit intermittent, control of error rate a plausible characteristic of the manual control problem. The specific logic which would control the switching between error and error-rate tracking is simply undetermined at this time.

Finally, let us note the interesting possibility that motion cues may be tracked in intermittent fashion just as nonvestibular error rate has been hypothesized to be here. Indeed, a very plausible theory for longitudinal pilot-induced oscillations has recently been offered by Smith[25] in which a pilot model was hypothesized which utilized switching between pitch attitude and normal acceleration tracking. While differing considerably in detail, Smith's model and the model discussed here are in consonance in treating the pilot as a single-channel processor in terms of visual (and/or aural or tactile) or vestibular stimuli and in assuming that error rates and accelerations are controlled independently, i.e., as variables in their own right.

In conclusion, a model of the human pilot has been offered which is an outgrowth of work reported in Refs. 9-11. Two features of the model distinguish it from other representations of the human pilot. First, proprioceptive information from the control stick or manipulator constitutes one of the major feedback paths in the model, providing feedback of vehicle output-rate due to control activity. Implicit in this feedback loop is a model of the vehicle dynamics valid in and beyond the region of crossover. Second, error-rate information is continuously derived and independently but intermittently controlled. An output injected remnant model was offered and qualitatively justified on the basis of providing a measure of the effect of inaccuracies such as time variations in the pilot's internal model of the controlled element dynamics. By varying the values of only five parameters, the data from experimental tracking tasks involving five different controlled element dynamics and one nonideal viewing condition were matched with model-generated describing functions and remnant power spectral densities. These model-generated results were obtained using a computational scheme which approximated the describing functions and remnant power spectral densities which would have been obtained if the model had been physically implemented and classical spectral techniques employed in measurement. The controlled element dynamics varied in terms of control difficulty from the nondemanding K and K/s dynamics to the $K/s(s-1)$ and $K/s(s^2 + 1.414s + 1)$ dynamics which approached the limits of manual control. It was indicated that the model characteristics which resulted from matching the data for the peripheral viewing experiments were qualitatively similar to measured pilot characteristics for a variety of tasks in which display quality was degraded.

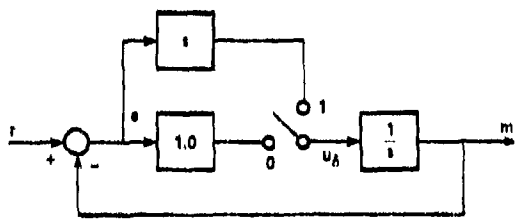
Appendix: Fourier Measurements in a Simple Switching System

Figure 8 is a simple model of a compensatory control system involving intermittent use of error-rate information. The input $r(t)$ consists of ten unity amplitude sinusoids with frequencies which are spaced in roughly uniform fashion between 0.2 and 14 rad/sec. Each sinusoid is repeated an integral number of times in the selected "run length" of 120 sec. The system of Fig. 8 was simulated on a digital computer. The switch operated in random fashion with the fraction of the run length spent in position 1 (P_1) equal to 0.2 in all cases. Three different minimum durations of switch closure were utilized: T_c = 0.3, 1.0 and 5.0 sec. These values encompassed minimum closure durations deemed acceptable for human tracking activity. The Fourier coefficients $C_{u_\delta}(j\omega_k)$ of the signal $u_\delta(t)$ were measured at each of the input frequencies ω_k with the switch in operation. In addition, the following calculation was performed:

$$[C_{u_\delta}(j\omega_k)]' = (1 - P_1)C_{u_\delta}(j\omega_k,0) + (P_1)C_{u_\delta}(j\omega_k,1) \quad (A1)$$

where $C_{u_\delta}(j\omega_k,i)$ indicates the coefficient measured with the switch in position i for the entire run. Table 2 shows the results of the simulation measurements in addition to the coefficients obtained by Eq. (A1) above. Also included are the $C_{u_\delta}(j\omega_k,i)$ values.

It was found that for the values of T_c used, the minimum duration of switch closure has no dominant effect upon the measured coefficients. Variations in the measured coefficients do occur with different values of T_c, however, these variations were found to be primarily due to the random switching activity itself. Consequently, the measured coefficients in the table have been obtained as the averages of the coefficients measured with the three values of T_c. As the table indicates, the truncated Taylor series of Eq. (A1) yields very



$$r(t) = \sum_{k=1}^{10} \sin(\omega_k t + \psi_k)$$

Fig. 8 Simplified switching model

Table 2 Fourier coefficient calculations for model of Fig. 8

Input Frequency	$C_{u_\delta}(j\omega_k)$ (measured)		$\left[C_{u_\delta}(j\omega_k)\right]'$		$C_{u_\delta}(j\omega_k,0)$ (measured)		$C_{u_\delta}(j\omega_k,1)$ (measured)	
ω_k (rad/sec)	Real	Imag	Real	Imag	Real	Imag	Real	Imag
.2094	-.16	-.45	-.17	-.41	-.20	-.45	-.053	-.25
.3142	.48	-.07	.42	-.085	.46	-.11	.25	.017
.5236	-.36	.16	-.38	.13	-.41	.16	-.25	-.026
.7330	-.005	.35	.041	.36	.078	.40	-.11	.23
1.152	-.28	.022	-.30	.038	-.32	.083	-.21	-.14
2.408	-.12	-.16	-.10	-.15	-.15	-.12	.072	-.24
3.874	-.09	-.072	-.10	-.069	-.12	-.024	-.015	-.25
5.178	.009	.087	.012	.086	.054	.039	-.134	.20
9.320	-.065	-.021	-.066	-.021	-.047	.026	-.142	-.21
14.35	.055	.021	.054	.023	.029	-.020	.156	.20

reasonable approximations to the measured coefficients when switching is occurring. It also indicates that the switching process can be adequately parameterised by the single quantity P_1, the probability of the switch being in position 1.

Obviously, this exercise with a single simple model was not meant to constitute a rigorous validation of the Fourier coefficient approximation technique utilised in the main body of this paper. Rather it was intended to offer quantitative evidence of the legitimacy of such an approximation technique when applied to a model which is generically similar to that investigated in the paper.

References

[1]McRuer, D. T., Graham, D., Krendel, E., and Riesener, W. Jr., "Human Pilot Dynamics in Compensatory Systems," Air Force Flight Dynamics Laboratory, AFFDL-TR-65-15, 1965.

[2]Kleinman, D. L., Baron, S., and Levison, W. H., "An Optimal Control Model of Human Response, Part I," Automatica, Vol. 6, 1970, pp. 357-369.

[3]Clement, W. F., McRuer, D. T., and Klein, R. H., "Systematic Manual Control Display Design," Guidance and Control Displays, AGARD CP-96, Oct. 1971, pp. 6-0 to 6-10.

[4]Hess, R. A., "Analytical Display Design for Flight Tasks Conducted Under Instrument Meteorological Conditions," IEEE Transactions on Systems, Man and Cybernetics, Vol. SMC-7, No. 6, June 1977, pp. 453-462.

[5]Anderson, R. O., "A New Approach to the Specification and Evaluation of Flying Qualities," Air Force Flight Dynamics Laboratory, AFFDL-TR-69-120, 1970.

[6]Hess, R. A., "Prediction of Pilot Opinion Ratings Using an Optimal Pilot Model," Human Factors, Vol. 19, No. 5, 1977, pp. 459-475.

[7]Ringland, R. E., Stapleford, R. L., and Magdaleno, R. E., "Motion Effects on an IFR Hovering Task - Analytical Predictions and Experimental Results," NASA CR-1933, 1971.

[8]Levison, W. H. and Junker, A. M., "A Model for the Pilot's Use of Motion Cues in Roll-Axis Tracking Tasks," Aerospace Medical Research Laboratory, AMRL-TR-77-40, 1977.

[9]Smith, R. H., "A Theory for Handling Qualities with Applications to MIL-F-8785B," Air Force Flight Dynamics Laboratory, AFFDL-75-119, 1976.

[10]Hess, R. A., "A Dual-Loop Model of the Human Controller," Journal of Guidance and Control, Vol. 1, No. 4, July-August 1978, pp. 254-260.

[11]Hess, R. A., "A Rationale for Human Operator Pulsive Control Behavior," Journal of Guidance and Control, Vol. 2, No. 3, May-June 1979, pp. 221-227.

[12]Kolers, R. A., Aspects of Motion Perception, Pergamon Press, New York, 1972, Chapter 3.

[13]McRuer, D. T., Hofmann, L. G., Jex, H. R., Moore, G. P., Phatak, A. V., Weir, D. H., and Wolkovitch, J., "New Approaches to Human Pilot/Vehicle Dynamic Analysis," Air Force Flight Dynamics Laboratory, AFFDL-TR-67-150, 1968.

[14]Magdaleno, R. E. and McRuer, D. T., "Experimental Validation and Analytical Elaboration for Models of the Pilot's Neuromuscular Subsystem in Tracking Tasks," NASA CR-1757, 1971.

[15]Granit, R., The Basis of Motor Control, Academic Press, London and New York, 1970.

[16]Kessel, C. and Wickens, C. D., "The Internal Model: A Study of the Relative Contribution of Proprioception and Visual Information to Failure Detection in Dynamic Systems," Proceedings of the Fourteenth Annual Conference on Manual Control, 1978, pp. 85-97.

[17]Levison, W. H., Baron, S., and Kleinman, D. L., "A Model for Human Controller Remnant," IEEE Transactions on Man-Machine Systems, Vol. MMS-10, December 1969, pp. 101-108.

[18]Skelton, R. E. and Likins, P. L., "Techniques of Modeling and Model Error Compensation in Linear Regulator Problems," in Control and Dynamic Systems, Vol. 14, Edited by C. T. Leondes, 1978, pp. 3-98.

[19]Levison, W. H., Elkind, J. I., and Ward, J. L., "Studies of Multivariable Manual Control Systems, A Model for Task Interference," NASA CR-1746, 1971.

[20]Levison, W. H., "The Effects of Display Gain and Signal Bandwidth on Human Controller Remnant," Aerospace Medical Research Laboratory, AMRL-TR-70-93, 1971.

[21]Pew, R. W., Duffendack, T. C. and Fensch, L. K., "Summary of Sine-Wave Tracking Studies," Proceedings of the Second Annual Conference on Manual Control, 1966, pp. 13-24.

[22]Jex, H. R., Allen, R. W., and Magdaleno, R. E., "Display Format Effects on Precision Tracking Performance, Describing Functions, and Remnant," Aerospace Medical Research Laboratory, AMRL-TR-71-63, 1971.

[23]Jex, H. R. and Allen, R. W., "Research on a New Human Dynamic Response Test Battery," Proceedings of the Sixth Annual Conference on Manual Control, 1970, pp. 743-777.

[24]Schmid, H. P. and Bekey, G. A., "Tactile Information Processing by Human Operators in Control Systems," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-8, No. 12, Dec. 1978, pp. 860-866.

[25]Smith, R. H., "A Theory for Longitudinal Short-Period Pilot Induced Oscillations," Air Force Flight Dynamics Laboratory, AFFDL-TR-77-57, 1977.

FITTS' LAW AND TARGET ACQUISITION

Richard J. Jagacinski
Ohio State University
Columbus, Ohio 43210

Daniel W. Repperger
Aerospace Medical Research Laboratory (AMRL/ME)
Wright-Patterson Air Force Base, Ohio 45433

Sharon L. Ward, Betty Glass, Martin S. Moran
Systems Research Laboratories, Inc.
2800 Indian Ripple Road
Dayton, Ohio 45440

ABSTRACT

Subjects used either a position or velocity control system to capture stationary and moving targets. For stationary targets, capture time was found to be linearly proportional to Fitts' Index of Difficulty, $\log_2(2A/W)$, for both types of control system. A represents the initial target displacement, and W represents target width. The incremental cost in capture time for an increase in the Index of Difficulty was considerably greater for the velocity control system. Phase-plane analyses of capture trajectories revealed a sequence of submovements. The first submovement was generally slower and more accurate than subsequent submovements. For moving targets, Fitts' Index of Difficulty was a good predictor of capture time with the velocity control system, but not with the position control system. An alternate Index of Difficulty measure is proposed which explicitly incorporates a velocity factor.

Fitts' Law (ref. 1) states that the time to move a stylus from a home position to a stationary target is proportional to an Index of Difficulty, $\log_2(2A/W)$. A is the distance from the home position to the center of the target, and W is the target width. A similar task to moving a physical stylus is to require a person to manipulate a control stick in order to move the displayed image of a cursor from a home position to a target region. In this latter situation it is possible to vary the relationship between the control stick manipulation and the resulting cursor movement by introducing system dynamics. For example, Sheridan and Farrell (refs. 2 & 3) and McGovern (ref. 4) investigated the relationship between speed and accuracy in performing discrete movements with remote manipulators. Jagacinski, Hartzell, Ward, and Bishop (ref. 5) recently found that Fitts' Law held for discrete movements performed with a velocity control system. However, the incremental cost in movement time for an increase in the Index of Difficulty was considerably greater than for a position control system. The first experiment in the present study was designed to examine the details of movement trajectories in this task in order to better understand the movement structure underlying Fitts' Law.

EXPERIMENT 1

Four subjects used a position control system and four subjects used a velocity control system to perform discrete movements from a home position to a stationary target. Subjects primarily used wrist rotation in order to manipulate a joystick. The cursor and target were displayed on an oscilloscope screen 1 m from the subjects. The display was "inside-out" so that movement of the joystick to the right resulted in the target moving to the left and the cursor remaining in the center of the oscilloscope screen. The display thus simulated the effect that would be obtained by centering a "viewing window" over an external stationary target. Gains of .418° and 3.0°/s of visual angle per 1° of control stick rotation were chosen for the position and velocity systems, respectively.

The cursor consisted of a single 18-cm vertical line, and the target consisted of a pair of 9-cm vertical lines that appeared randomly to the right or left. The targets were generated from a factorial combination of three amplitudes (A = 48, 84, and 147 mm), three target widths (W = 3.00, 5.25, and 9.20 mm), and two directions (right and left). Two criteria were used for defining target capture. A position criterion required that a smoothed estimate of the cursor position remain over the target for 350 ms. A position-plus-steadiness criterion additionally required that a smoothed estimate of cursor velocity remain less than .5°/s over the 350 ms interval. When the criterion that was in effect for a given trial was met, the trial was over and the cursor and target disappeared from the screen. Subjects were instructed to try to get the target to the center of the screen and make the screen go blank as quickly as possible. The criterion for ending a trial alternated across days in an ABBA manner between the position and position-plus-steadiness criteria.

Movement time was measured from when the subject began to move the control stick with a velocity of 2°/s until the beginning of the 350 ms capture criterion. Subjects received 216 trials per day in which the targets were presented in random order. Subjects approached asymptotic performance by Days 9-12, and the average movement times for these days are shown in Figure 1. The slopes were considerably steeper for the velocity control system (278 ms/bit for the position criterion; 240 ms/bit for the position-plus-steadiness criterion) relative to the position control system (169 ms/bit for the position criterion; 175 ms/bit for the position-plus-steadiness criterion). Secondly, the intercepts were considerably higher for the position-plus-steadiness criterion for both control systems.

In order to relate these effects to the microstructure of the movements, phase-plane trajectories were examined from digital magnetic tape recordings of the smoothed cursor position and velocity sampled every 5 ms. Examples of phase-plane trajectories for subjects with the lowest and highest peak velocities for each of the two control systems are shown in Figure 2. The trajectories generally do not resemble the straight line, parabolic, or spiral shapes that one would expect from linear first or second-order continuous systems. Rather, the terminal sections of these movements usually consist of a succession of somewhat irregular episodes of acceleration and deceleration, and these episodes may be termed "submovements."

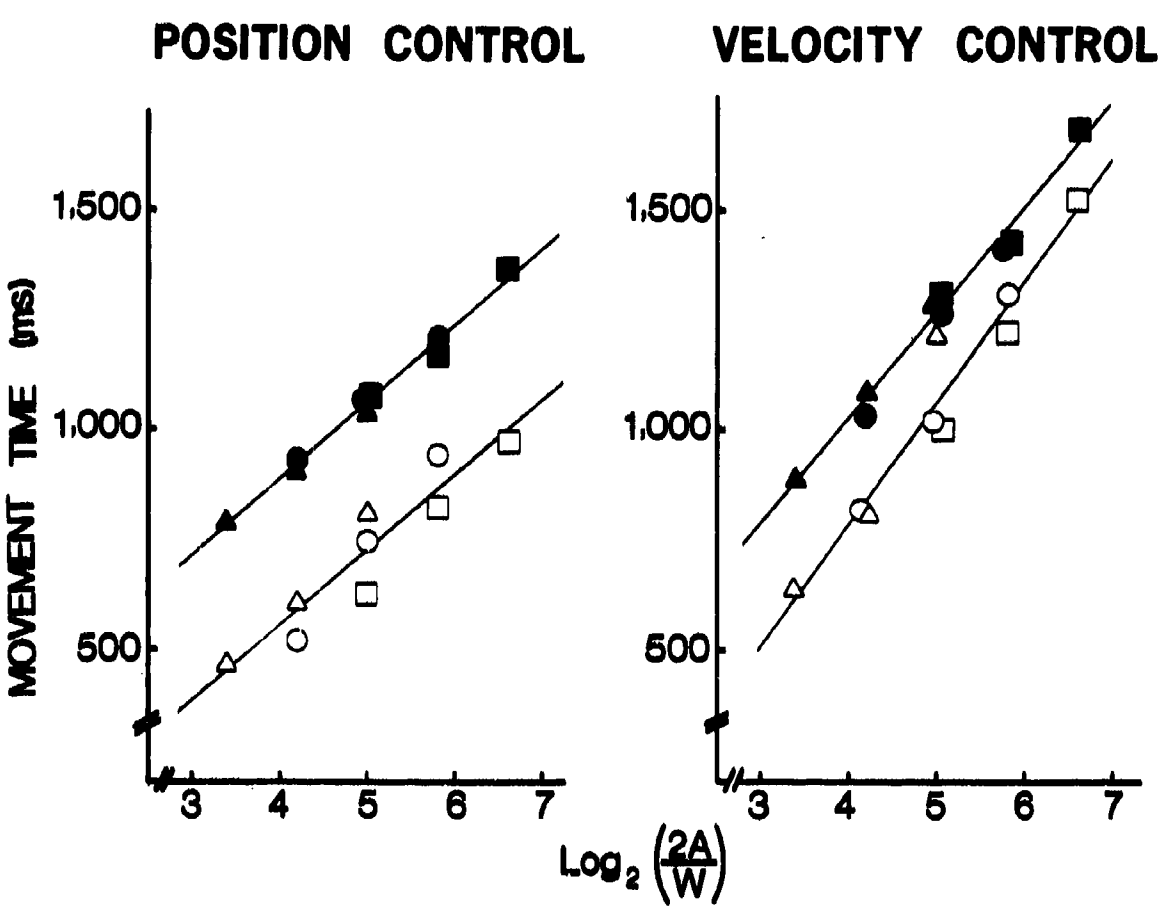


Figure 1. Average movement times for four subjects using a position control and four subjects using a velocity control. Open symbols indicate a position criterion and solid symbols indicate a position-plus-steadiness criterion. Triangles, circles, and squares respectively indicate the shortest, intermediate, and longest movement amplitudes, A.

In order to quantify these episodes, the ending of a submovement was defined as occurring when 1) the velocity changed sign indicating a shift in movement direction, or 2) when the velocity did not change sign, but the trajectory decelerated and then reaccelerated toward the target. The beginning of a submovement was taken as the end of each previous submovement. Two types of episodes were excluded from this analysis: 1) submovements for which the relative accuracy (absolute value of the ratio of final distance from the target center to initial distance from the target center) was greater than 1.0, and were therefore counterproductuve; 2) submovements that lasted less than 70 ms. This latter criterion was necessary in order to exclude system noise that would otherwise have been mistaken for submovements generated by the subjects. The former criterion eliminated submovements away from the target and occasional gross overshoots.

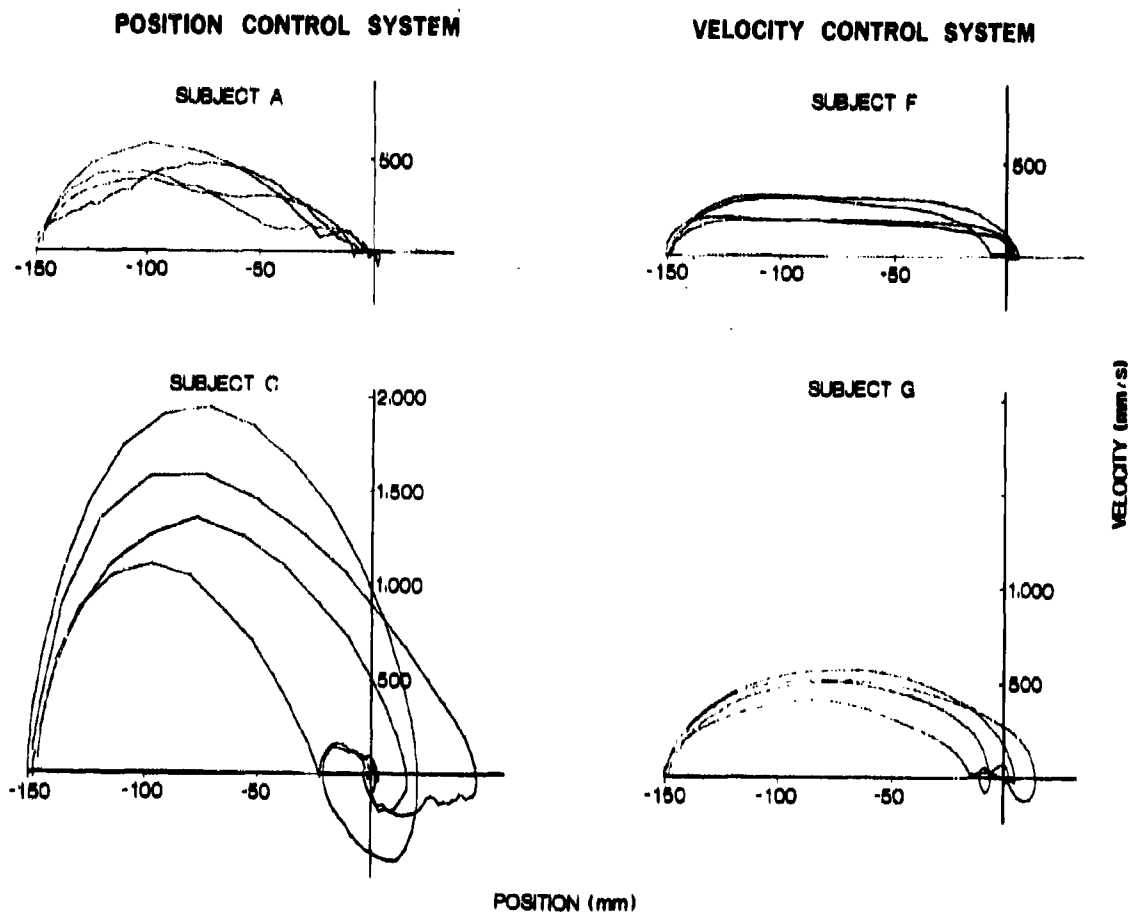


Figure 2. Phase-plane trajectories of four movements (A = 147 mm, W = 3.00 mm) for each of four subjects using a position criterion.

Median duration and accuracy of the first submovements, as well as second, third, and fourth submovements when they were present, are displayed in Figures 3 and 4 for each subject's last day of performance with each of the two ending criteria. With only one exception (Subject C, position criterion), the first submovement medians were both slower and more accurate than the medians of subsequent submovements. This dichotomy is extremely marked for the velocity control system. There is also a strong correlation between the duration and log-accuracy of second, third, and fourth submovements for the position control system (r = .89 for the position criterion; r = .98 for the position-plus-steadiness criterion). This strong correlation is not present for the velocity control system (r = .38 for the position criterion; r = .45 for the position-plus-steadiness criterion).

Averaging the median durations and log-accuracies across subjects, the ratio of mean duration to mean log-accuracy for second, third, and fourth submovements was 134 ms/bit (position criterion) and 139 ms/bit (position-plus-steadiness criterion) for the position control system, and 277 ms/bit (position criterion) and 247 ms/bit (position-plus steadiness criterion) for the velocity control system. These ratios are within three percent of the obtained MT slopes for the velocity control system. For the position control system, these values

are about 80% of the obtained MT (movement time) slopes in Figure 1.

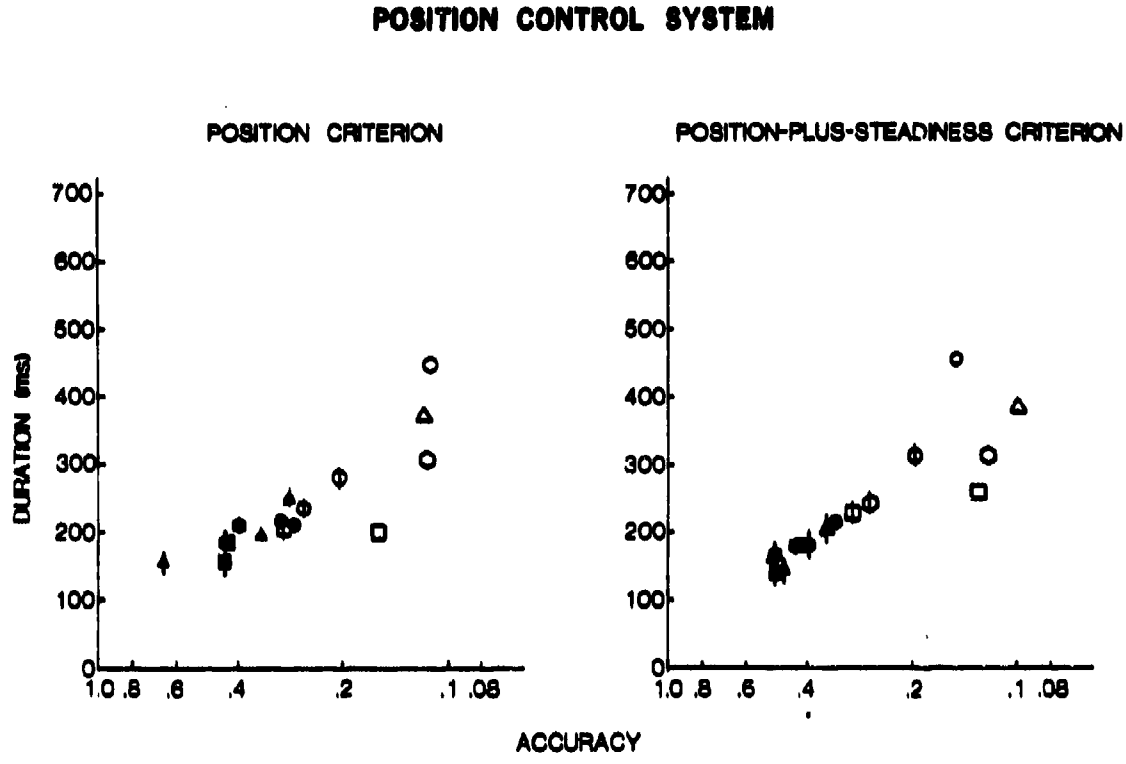


Figure 3. Median accuracy and duration of submovements of Subjects A (circles), B (hexagons), C (squares), and D (triangles), who all used a position control system. First submovements are open symbols; second submovements are open symbols with a vertical line; third submovements are solid symbols; fourth submovements are solid symbols with a vertical line.

The ending criteria seem to have had little effect on the overall medians of second, third, and fourth submovements, and the effects on the overall medians of first submovements were not consistent across subjects. The effect of the position-plus-steadiness criterion in raising movement times appears to be primarily due to the presence of additional submovements rather than to a change in submovement structure. The average of the individual subject's median number of submovements to the beginning of the 350 ms capture criterion interval was 2.1 (position criterion) and 3.1 (position-plus-steadiness criterion) for the position control system, and 2.5 (position criterion) and 3.0 (position-plus-steadiness criterion) for the velocity control system. Multiplying the number of submovements by the average of second, third, and fourth submovement durations and subtracting the results for the two ending criteria predicts an increase in movement time with the position-plus-steadiness criterion of 180 ms with the position control and 128 ms with the velocity control. These values are, respectively, 53% and 63% of the observed differences in movement times in Figure 1.

VELOCITY CONTROL SYSTEM

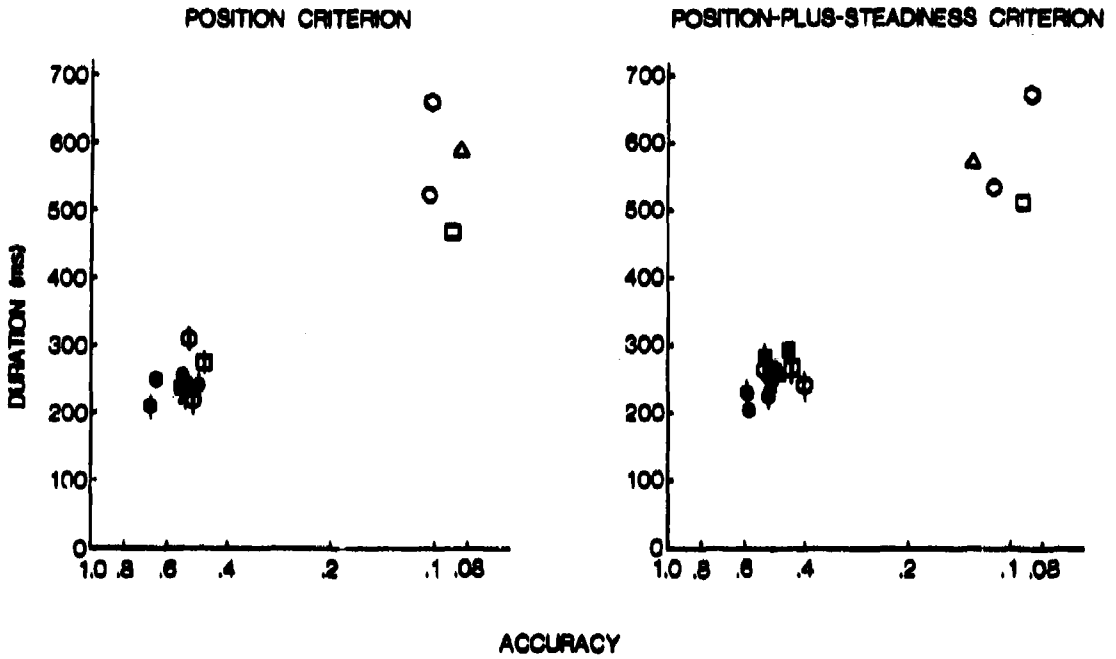


Figure 4. Median accuracy and duration of submovements of Subjects E (circles), F (hexagons), G (squares), and H (triangles), who all used a velocity control system. First submovements are open symbols; second submovements are open symbols with a vertical line; third submovements are solid symbols; fourth submovements are solid symbols with a vertical line.

EXPERIMENT 2

One way of characterizing the subjects' task in Experiment 1 is as a step-tracking task in which the target width is explicitly indicated. Experiment 2 examined an analogous constant velocity ramp tracking task in which the target width was also explicitly indicated. The basic question of interest was whether Fitts' Law would generalize to moving targets.

The control systems and display were the same as in Experiment 1. Targets were generated from a factorial combination of three amplitudes ($\underline{A}$ = 48, 84, and 147 mm), three targets widths ($\underline{W}$ = 5.25, 9.20, and 16.10 mm), three velocities ($\underline{V}$ = 0, 30.96, and 54.77 mm/s), and two directions (right and left). In terms of degrees of visual angle, these values respectively correspond to: $\underline{A}$ = 2.75°, 4.81°, and 8.42°; $\underline{W}$ = .30°, .53°, and .92°; $\underline{V}$ = 0°/s, 1.77°/s, and 3.14°/s. Targets with non-zero velocities were always moving toward the center of the oscilloscope screen at the beginning of a trial.

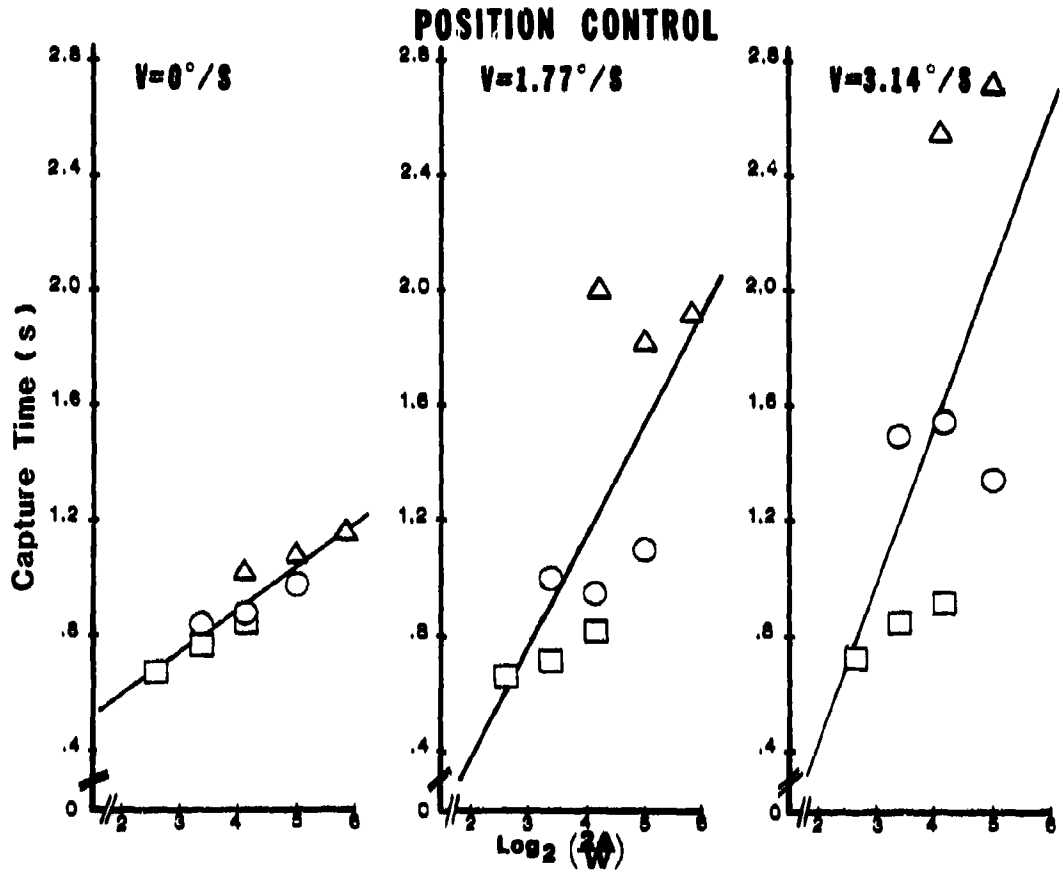


Figure 5. Average capture times vs. Fitts' Index of Difficulty for four subjects using a position control system. Triangles, circles, and squares respectively represent the smallest, intermediate, and largest target width.

The basic measure of performance in this task was capture time, the interval from the appearance of the target until the beginning of the 350 ms capture criterion. (Capture time is the sum of reaction time plus movement time.) Only the position ending criterion was used in this experiment. Subjects received 216 trials per day in which the targets were presented in random order. By Sessions 10-13, subjects approached asymptotic performance, and average capture times for four subjects using a position control system and four subjects using a velocity control system are shown in Figures 5 and 6. The data point for the position control system with the target having the highest velocity, largest initial displacement, and narrowest width is omitted because subjects failed to capture it within a 4-s time limit on 50% of the trials it appeared.

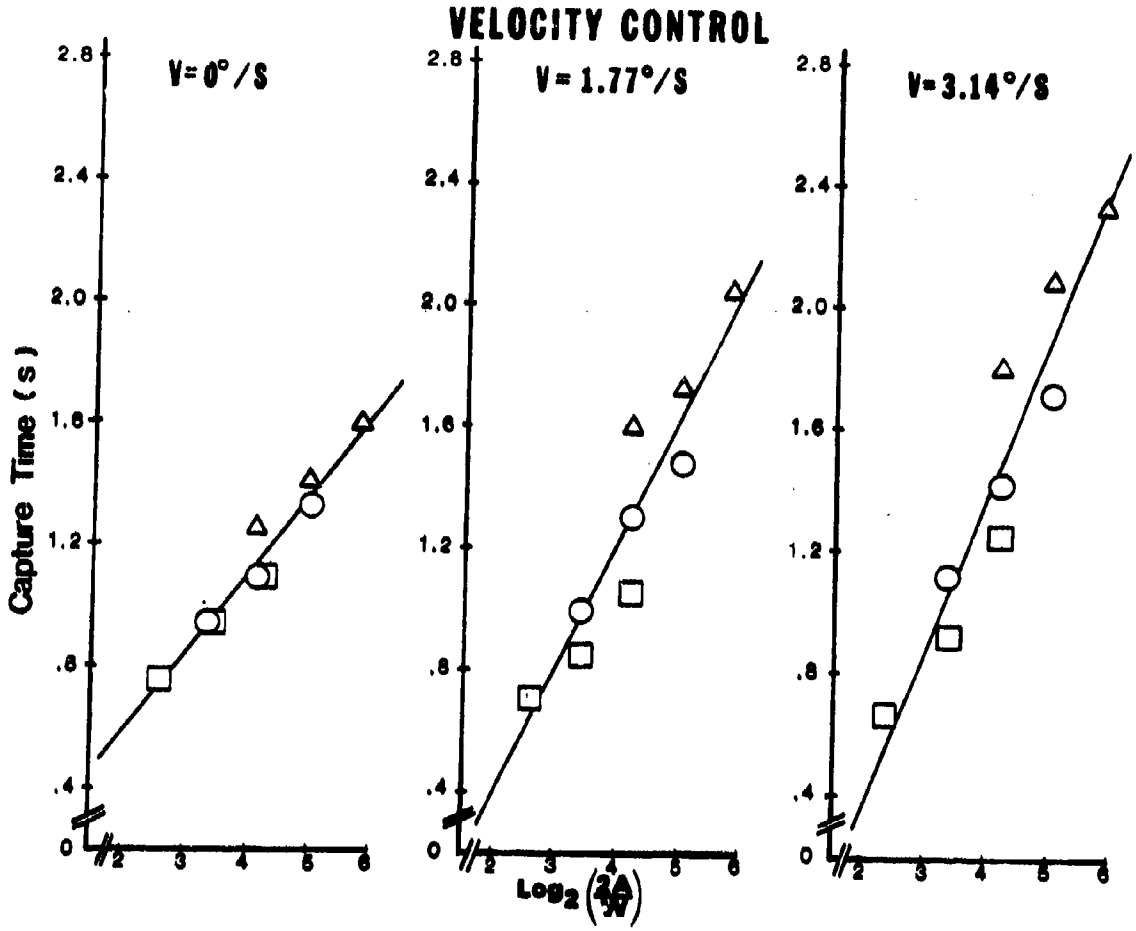


Figure 6. Average capture times vs. Fitts' Index of Difficulty for four subjects using a velocity control system. Triangles, circles and squares respectively represent the smallest, intermediate, and largest target width.

For the stationary targets ($\underline{V}$ = 0°/s), the capture times for both the position and velocity control systems were well approximated by a linear function of $\log_2(2\underline{A}/\underline{W})$. The slope for the velocity control (264 ms/bit) was considerably steeper than for the position control (147 ms/bit), which agreed well with the results of Experiment 1. For the moving targets ($\underline{V}$ = 1.77°/s and 3.14°/s) a linear function of $\log_2(2\underline{A}/\underline{W})$ provided a reasonably good approximation to the data for the velocity control system, with the slopes becoming steeper for the non-zero velocities. However, for the position control system $\log_2(2\underline{A}/\underline{W})$ was not a very good predictor of capture time for the moving targets. For this control system, the increase in velocity had little effect on capture time for the widest target, but very marked effects for the two narrower targets. Furthermore, the effect of the initial target displacement, $\underline{A}$, did not appear to be similarly increased by the increasing velocity.

One possibility for a new Index of Difficulty that will describe these effects is the equation:

$$CT = c + dA + e(V + 1)\left(\frac{1}{W} - 1\right) \qquad (1)$$

where CT is capture time in seconds, A, V, and W are specified in degrees of visual angle, and c, d, and e are regression coefficients. This equation was one of about a dozen candidates that the authors tested, and it has a number of qualitative structural characteristics that correspond well to features of the data. As in the expression $\log_2(2A/W)$, this new equation also predicts longer capture times as A becomes larger and W becomes smaller. The lengthening of capture times arising from the strong interaction of target width and velocity is reflected in the multiplication of the (V + 1) and $(\frac{1}{W} - 1)$ terms. The constant "1" is added to V so that even when V = 0°/s the effect of W on capture time will be maintained in this representation. The constant "1" is subtracted from 1/W so that the interaction with velocity will be minimal for the widest target (1/W = 1.08/°). The subtraction of "1" may therefore be considered as a fourth fitting parameter, although it was not varied between the representations for the position and velocity control systems.

The success of this new representation for the Index of Difficulty is displayed in Figure 7. The variable A was not significant for the position control system, and so was omitted from the predictor equation. The standard error was around .1 s for each control system, which is relatively small in comparison with the large range of capture times. When an attempt was made to use Fitts' Index of Difficulty to fit the capture times for all three velocities simultaneously, the standard error was approximately .2 s for the velocity control system and .5 s for the position control system.

DISCUSSION

In summary, it appears that Fitts' Index of Difficulty is an excellent predictor of capture times for stationary targets with either a position or velocity control system. In attempting to understand the movement processes underlying this relationship, researchers have formulated two classes of models. One class postulates a unitary pattern of convergence toward the target; the other class postulates a dichotomous convergence pattern. The analysis of submovement accuracy and duration in Experiment 1 indicates that the first submovement is quite different from subsequent submovements. This finding rejects models of the first class such as the simple discrete submovement models of Crossman and Goodeve (ref. 6) or a continuous second-order underdamped linear system model of discrete movement (ref. 7). Both of these models posit constant accuracy and constant duration across submovements. Keele (ref. 8) proposed a model very similar to that of Crossman and Goodeve (ref. 7), the major difference being that the first submovement was assumed to be of shorter duration than (but of equal accuracy to) subsequent submovements. This model can also be rejected for the present data. Other investigators such as Welford (ref. 9) who have posited a strong dichotomy between first and subsequent submovements have generally hypothesized that the first submovement is open-loop and subsequent submovements are under feedback control. Although

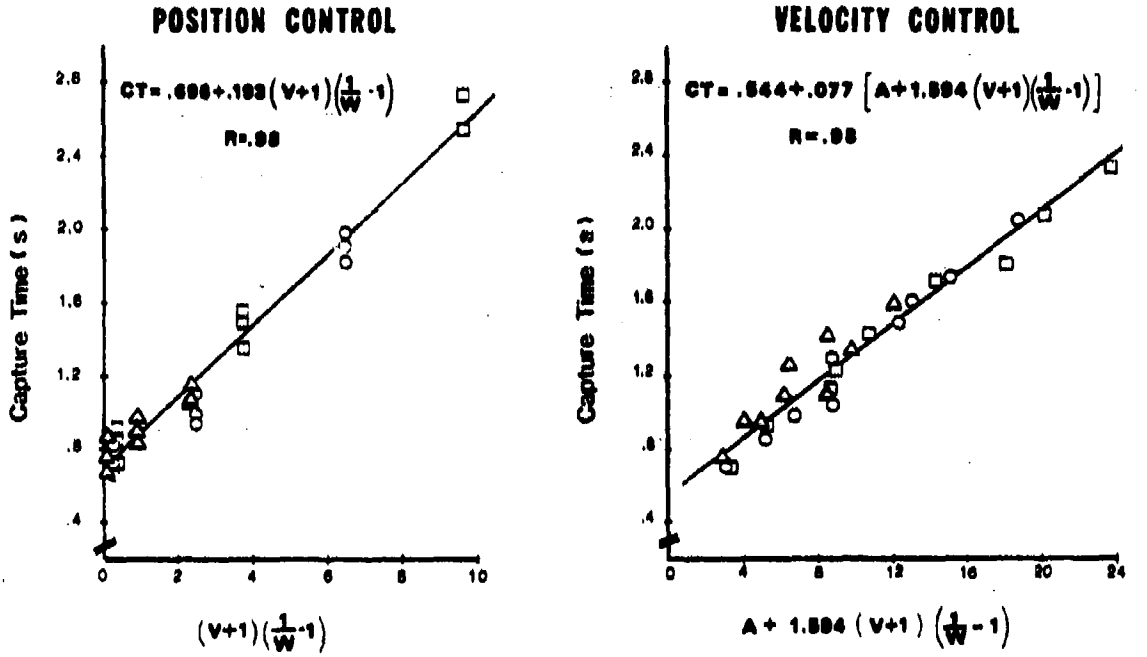


Figure 7. Average capture times replotted vs. modified Indices of Difficulty. Triangles, circles, and squares respectively represent target velocities of 0°/s, 1.77°/s, and 3.14°/s.

the present data are insufficient to establish conclusively the use of open or closed-loop control, the relatively long durations of the first submovements suggest that they may not be open-loop.

In general, the overall duration and accuracy of a movement will be equal to the sum of the submovement durations and the product of submovement accuracies. The accuracy of a submovement is taken to be the absolute value of final distance from the target center to initial distance from the target center. Accuracies must be logarithmically scaled to represent their concatenation additively. In that $2A/W$ or equivalently $A/(W/2)$ is the inverse of the accuracy of the overall movement, Fitts' Index of Difficulty provides such a log-accuracy scale. The time course of a movement can thus be traced out in the movement time vs. log-accuracy space of Figure 1, each submovement adding an increment of time and adding an increment of log-accuracy. If one ignores speed-accuracy variations among the first submovements for different targets, then a sufficient condition for generating a set of movement times obeying Fitts' Law is that the ratio of duration to log-accuracy be constant for second and subsequent submovements. The submovements for the velocity control system approximately satisfy this condition and the ratio of mean duration to mean log-accuracy of the second, third, and fourth submovements gives an excellent estimate of the overall movement time slopes. This simple means of estimating the overall movement time slope is somewhat less accurate for the position control system.

For moving targets, Fitts' Index of Difficulty is a good predictor of capture times for the velocity control system, but not for the position control system. The new Index of Difficulty which represents the interaction of velocity and target width is a better predictor. However, this new predictor is still merely an approximation to the pattern of capture times, and it may be possible to design even better predictor equations. Submovement analyses have not been conducted on the capture trajectories for the moving targets, so no additional data are available as to the basis of the interaction of width and velocity.

The authors are hopeful that the Indices of Difficulty discussed in the present paper will prove useful in modeling tasks that are composed of a mixture of target acquisition and continuous tracking tasks. However, additional research may be needed on the interfacing of the equations predicting the achievement of target capture and continuous tracking models describing subsequent following performance. The philosophy behind such an overall modeling effort would be that the correction of large tracking errors and small tracking errors require separate descriptions rather than being approximated by a single control strategy. This dichotomy may be analogous to the dichotomy observed between first and subsequent submovements.

REFERENCES

1. Fitts, P. M., & Peterson, J. R. Information capacity of discrete motor responses. Journal of Experimental Psychology, 1964, 67, 103-112.

2. Sheridan, T. B., & Ferrell, W. R. Remote manipulative control with transimssion delay. IEEE Transactions on Human Factors in Electronics, 1963, HFE-4, 25-29.

3. Ferrell, W. R. Remote manipulation with transmission delay. IEEE Transactions on Human Factors in Electronics, 1965, HFE-6, 24-32.

4. McGovern, D. E. Factors affecting control allocation for augmented remote manipulators. Unpublished Ph.D. dissertation, Stanford University, 1975.

5. Jagacinski, R. J., Hartzell, E. J., Ward, S., & Bishop, K. Fitts' Law as a function of system dynamics and target uncertainty. Journal of Motor Behavior, 1978, 10, 123-131.

6. Crossman, E. R. F. W., & Goodeve, P. J. Feedback control of hand-movement and Fitts' Law. Experimental Psychology Society Meeting, July 1963.

7. Langolf, G. D., Chaffin, D. B., & Foulke, J. A. An investigation of Fitts' Law using a wide range of movement amplitudes. Journal of Motor Behavior, 1976, 8, 113-128.

8. Keele, S. W. Movement control in skilled motor performance. Psychological Bulletin, 1968, 70, 387-403.

9. Welford, A. T. Fundamentals of skill. London: Methuen, 1968.

ACKNOWLEDGEMENTS

The present work was conducted under Air Force Office of Scientific Research Grant 77-3288 by Ohio State University and personnel of Systems Research Laboratories, Inc., under contract F33615-C-76-5001, jointly with the Aerospace Medical Research Laboratory, Aerospace Medical Division, Air Force Systems Command, Wright-Patterson Air Force Base, Ohio, under project 2312V601. The authors wish especially to thank Jim Hartzell, George Valentino, Carroll Day, Kaile Bishop, Robert McIntyre, Neil Rancour, Jennifer Zingg, Paul Isaac, and David Greene for their assistance in executing and analyzing the present experiment. Reprints of this article are identified by the Aerospace Medical Research Laboratory as AMRL-TR-79-26. Further reproduction is authorized to satisfy the needs of the U.S. Government.

MULTI-LEVEL AND MULTI-TASK SUPERVISORY CONTROL

Dana R. Yoerger, Bahman Daryarian and Thomas B. Sheridan
Massachusetts Institute of Technology, Cambridge MA 02139

ABSTRACT

This paper reviews recent experimental research on a NASA program concerned with aircraft pilots as supervisory controllers. In one study a simulated YC-15 is flown over a course changing in altitude and position with an without turbulence. Four different control modes are compared: (1) pitch rate, roll rate and horizontal thrust; (2) pitch, roll and horizontal velocity set points; (3) vertical speed, heading and velocity set points; (4) altitude, x-y position and velocity set points. A variety of performance measures including error control activity and subjective workload rating suggest that (3) is better than (2) and (2) than (1), but (4) is not better than (3). In considering transient "switch-down" of modes in the event of failure, (3) appears best to switch from, and (1) to switch to.

In a second experiment, a subject is given a variety of tasks to do having varying arrival times, importances, durations, distances from a deadline, and switching delays. He must allocate his one-channel of control to these tasks, sometimes ignoring some tasks to do the most important ones. We have formulated some hypothesis about how mental workload bears a non-monotonic relation to "task loading" in such a situation (after the Ph.D research of our colleague, Tulga) and are testing this hypothesis using a multi-dimensional scaling technique.

4

SESSION 4: CONTROL AND MANIPULATORS

Moderator: Dr. Ronald A. Hess, NASA Ames Research Center

HOW TO TALK TO A ROBOT MANIPULATOR

Thurston, L. Brooks and Thomas B. Sheridan
Massachusetts Institute of Technology, Cambridge MA 02139

ABSTRACT

In the MIT Man Machine Systems Lab. we have developed a seven-degree-of-freedom remote manipulator capable of being controlled in a range of modes. Conventional modes include: (1) force-reflecting master slave; (2) variable rate joystick; (3) position within a spatial envelope and rate outside that envelope. Simple supervisory modes include: (4) continuous trajectory teach-play back, taught by any of (1) (2) (3); (5) point-to-point teach - playback; (6) smoothed point-to-point teach - playback. "Smart" supervisory modes include: (7) concentration of previously defined sub-routines (8) branching based on externally sensed force; (9) teach in one manipulator configuration, initialize and playback with an arbitrarily different manipulator configuration. Experiments comparing performance in these modes independently and in combination are reviewed. Applications to deep ocean tasks will be discussed.

A TACTUAL AID FOR THE VISUAL APPROACH *

Sheryl L. Chappell **
Richard D. Gilson

The Ohio State University
Department of Aviation
P. O. Box 3022
Columbus, Ohio 43210

Abstract

The subtle visual approach information from the runway scene was supplemented in this study by heads-up tactual or visual displays of airspeed and glide path error. It was anticipated that the tactual presentation would greatly enhance novice pilot's learning of the relationships between the visual transformations in the runway environment and aircraft control during the visual approach. Novice pilots flew simulated visual approaches to landing in a ground based trainer. The experimental tasks during training and subsequent testing were control of both airspeed (pitch) and glide angle (throttle). Performance was measured during the initial training and during test approaches (where no displays were present) following each training session. The subjects in the display groups received one of the four combinations of visual and tactual displays for pitch and throttle, with no verbal assistance. Therefore, an additional control group received only verbal assistance (no displays) during training and no such assistance during testing.

Results.During the training periods, performance with the visual display of pitch information was significantly better than performance with the tactual pitch display. Throttle error did not vary for the two groups during the training periods.

In contrast to the results obtained during training, the group who received tactual pitch error information performed both pitch and throttle control tasks significantly better than the visual group without the presence of any displays. Testing for the group who had verbal instruction showed that their performance either equaled (for pitch) or exceeded (for throttle) that of the display groups.

* This work was accomplished under contract (# NAS 2-8954) from the National Aeronautics and Space Administration, Ames Research Center, Moffett Field, California 94035.

**Now affiliated with the Battelle Memorial Institute, Columbus, Ohio.

Conclusions. Under the conditions of this experiment, visual perceptual learning of the approach to landing task is more greatly facilitated by training with heads-up tactual rather than visual displays of the same information. Secondly, ample verbal assistance by itself yields a training advantage over either display. Verbal instruction together with the tactual display of control information during the approach, appears to be the best combination of techniques tested to improve the conventional training for the visual approach to landing.

Introduction

The objective of the study was to examine the sense of touch as a means of transmitting control information to aid the perceptual learning for the visual approach to landing.

Information for the visual approach comes from subtle changes in the runway scene and the cockpit instruments. It is known that learning to utilize the outside visual information is difficult for most pilots. Past research has shown that using other senses for information can reduce visual workload (Kahneman, 1973); this study was designed to explore the possibility of providing tactual control information to aid the learning of the subtle visual cues. Heads-up tactual and visual displays were compared as training aids.

The visual approach to landing has two major task components, control of the aircraft's airspeed and its glide angle to touchdown. Information from outside the cockpit for airspeed control comes mainly from the attitude of the aircraft's nose relative to the horizon. Airspeed corrections are made by altering the pitch with fore/aft movement of the control yoke. Outside information for the glide angle is more subtle. Pilots use the apparent shape of the runway as it changes from a rectangle to a trapezoid during the progression of the approach. Also, the projected point of touchdown will remain stationary in the center of an expanding optic array (Gibson, 1955). Glide angle corrections are primarily made by fore/aft movement of the throttle.

Precise tactual feedback of control errors does not visually interfere with outside visual cues. More importantly, tactual control information may facilitate perceptual learning of the use of these cues. The relationship among information received from the runway environment, airspeed errors, and glide angle errors may be more readily learned.

Methodology

Apparatus. The aircraft simulator was a Singer GAT-1 motion-based trainer with an interactive visual simulation of the runway. The appearance of the runway changed when the aircraft approached, as a function of the angle and distance along the glide path.

Heads-up tactual and visual displays presented airspeed (pitch) and glide angle (throttle) errors. The tactual display unit has a movable servo-controlled slide embedded in the control handle (Figure 1). When the slide is in the flush position no error exists.

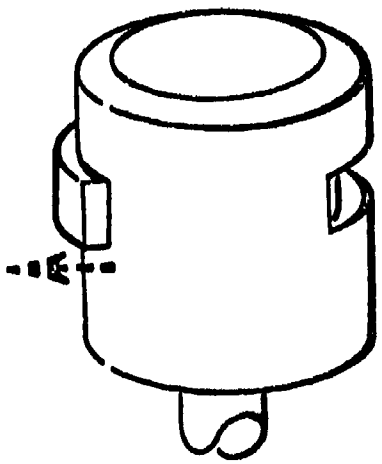


FIGURE 1

Kinesthetic-tactual display depicting a servo-controlled slide embedded in a control stick.

The slide is driven in the direction of the error. Control response in that direction corrects the error and therefore, the slide returns to the neutral position. Both kinesthesia and touch are used to detect the protrusion or recession of the slide, relative to the control handle. This display has been used successfully in many vehicle control applications, including simulated helicopter control displays in the cyclic and collective, (Gilson and Ventola, 1976) and its original application for car following (Fenton and Montano, 1968). The simulator used in this study was equipped with a tactual display for airspeed in the control yoke, and glide angle error in the throttle.

The heads-up visual displays consisted of two horizontal lines projected on the runway scene. A deviation from the horizon for the airspeed line and from the runway touchdown point, for the glide angle display indicated error. The displays were tracked (in a compensatory manner) in the same fashion as the tactual displays.

Subjects. The subjects were novice pilots who volunteered one hour a day for four days. The subjects were randomly assigned to five experimental groups. Group 1 received visual pitch and throttle displays; Group 2 received tactual displays and Groups 3 and 4 had one visual and one tactual display. The fifth or control group received verbal instructions from the experimenter.

Procedure. On the first day, approach information and control usage were explained to the subjects. Only pitch and throttle controls were used, roll and yaw were locked. Pitch was continually emphasized as the primary task. Initially, two demonstration approaches were flown jointly by the experimenter and subject with reference only to the outside runway simulation. Then the subject flew four approaches, often landing short of the runway. On the following days there were four experimental sessions, each consisting of 16 practice approaches with the displays and four test approaches without displays or verbal instruction. Data were recorded from the last four practice runs and all four of the test approaches.

Results

Two performance measures were analyzed. Airspeed and glide angle performance were measured by the pitch and throttle integral errors, respectively.

Pretraining. Performance on the first four approaches before the displays were introduced did not show a significant difference among groups before training. This result was expected due to the random assignment of subjects to groups.

Practice. The two groups with the visual pitch display had significantly less pitch error while the displays were present (Figure 2).

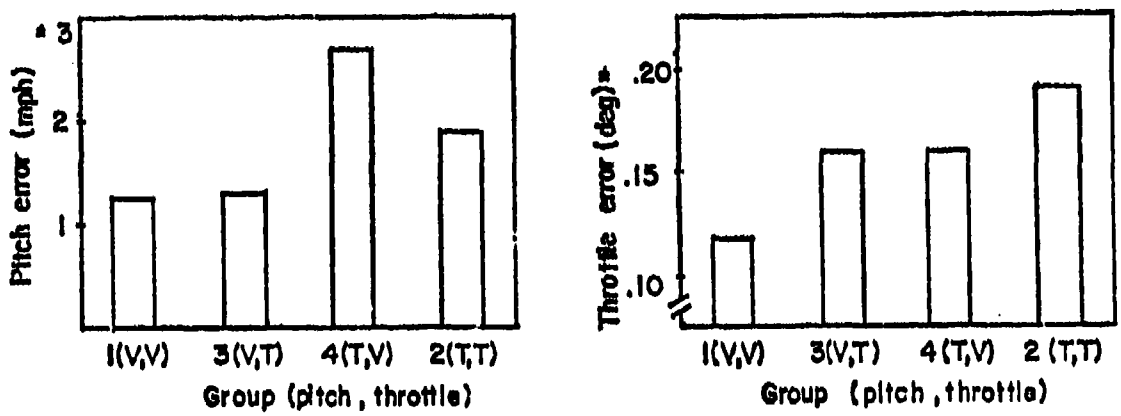


FIGURE 2

Group effects for visual (V) and tactual (T) displays; performance with the displays.

This was expected since the tactual displays had no velocity quickening. Previous studies had shown velocity quickening of the tactually displayed error was necessary to yield the same level of performance as with a visual display (Jagacinski, Miller, Gilson & Ault, 1977). Throttle performance with the displays showed no significant group effects.

Testing. In contrast, during test trials without the displays, subjects having practiced with tactually displayed pitch error had better performance than the visual group. This effect was significant for both pitch and throttle error (Figure 3).

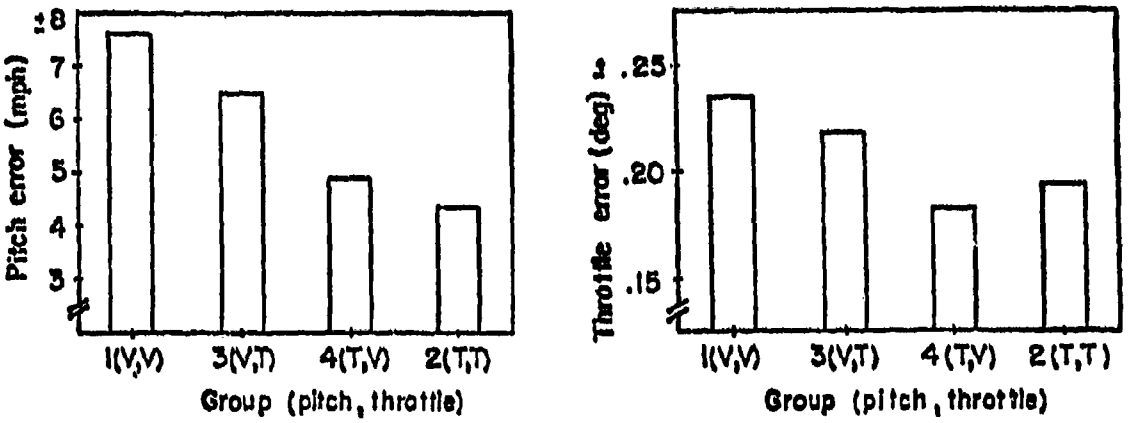


FIGURE 3

Group effects for visual (V) and tactual (T) displays; performance without the displays.

These were again the error levels for performance in the transfer condition when approaches were made with only outside runway information. The performance of the group with verbal instruction either equalled (for pitch error) or significantly exceeded (for throttle control) both the display groups.

Learning Effects. An examination of performance data without the displays across test periods showed that learning occurred. This effect was significant for the combined data from all groups for pitch and throttle error (Figure 4).

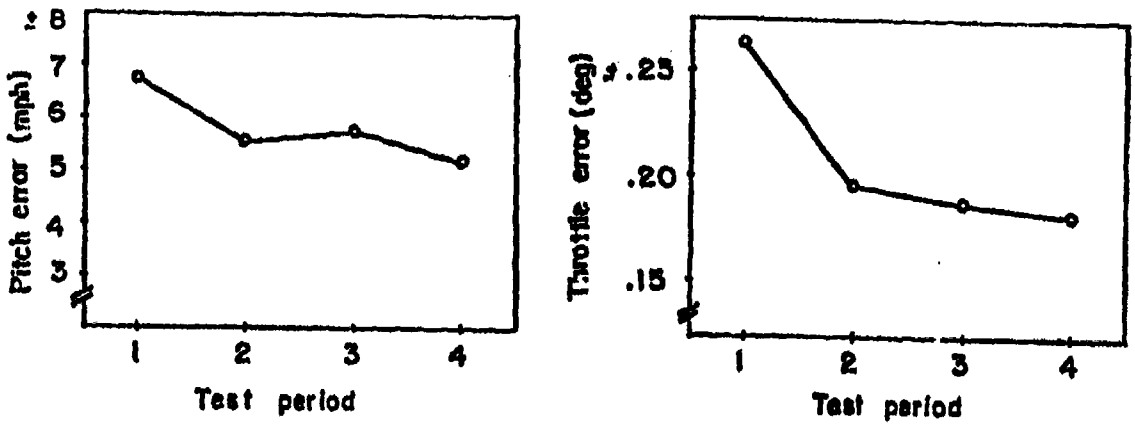


FIGURE 4

Test period effects for all subjects' performance without the displays

Learning effects while using the displays were significant only for throttle error (Figure 5).

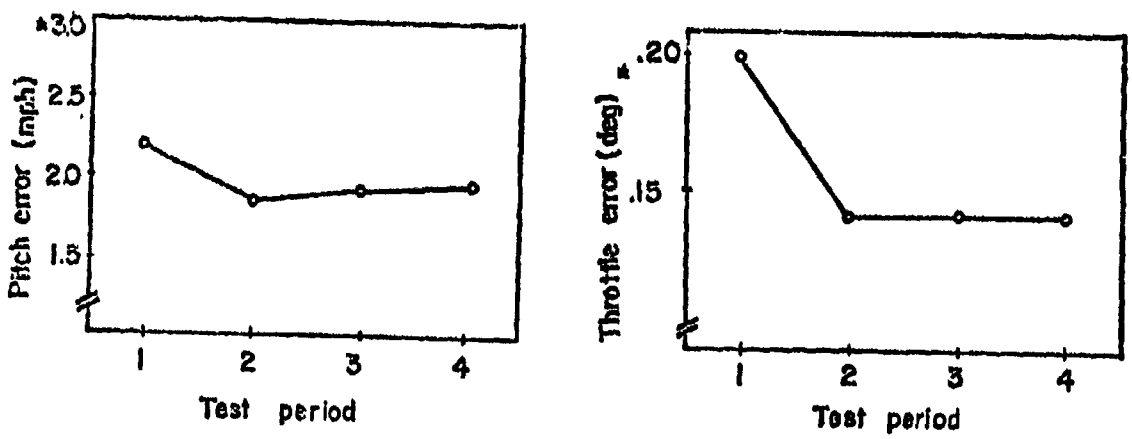


FIGURE 5

Test period effects for all subjects' performance with the displays.

Conclusions

In summary, tactual training enhanced the learning for use of outside information to make a visual approach to landing as did verbal instruction. Both teaching techniques reduce visual workload and allow the student to pay more attention to the runway environment. Verbal instruction was superior because it is not confined to control error and can encompass e.g. anticipation of needed response and control interactions. However, in flight these instructions may interfere with radio communications and attention to such things as engine sounds. Therefore, a combination of precise control information presented tactually and verbal coaching may improve teaching of the difficult task, the visual approach to landing.

References

Fenton, R.E. & Montano, W.B., An Intervehicular Spacing Display for Improved Car-Following Performance IEEE Transactions on Man-Machine Systems, June 1968, MMS-9(2).

Gibson, J.J., The Optical Expansion Pattern inAerial Locomotion., The American Journal of Psychology, September 1955, LXV III, 480-484.

Gilson, R.D., & Ventola, R.W., Evaluation of a Tactual Display Device with Adapting Primary and Secondary Task Performance Methods, (Final Report, Contract AMRDL CR-76-1), Moffett Field, CA.: United States Army, Air Mobility Research and Development Laboratory, Ames Research Center, February 1976.

Kahneman, D. Attention and Effort, Englewood Cliffs, N.J.: Prentice Hall, 1973.

EVALUATION OF PROXIMITY SENSOR AIDED GRASP CONTROL FOR SHUTTLE RMS

A. K. Bejczy
Member of the Technical Staff
Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91103

J. W. Brown
Aerospace Engineer

J. L. Lewis
Section Head
Crew Station Design Section
Lyndon B. Johnson Space Center
Houston, Texas 77058

SUMMARY

A proximity sensor system has been developed at the Jet Propulsion Laboratory (JPL) and integrated with a four-claw end effector of Johnson Space Center (JSC). The sensor system has been designed to aid the operator to find the proper terminal range and pitch and yaw alignments of the four-claw end effector on a 15.2-m (50-ft) long manipulator relative to the grapple fixture of a large payload. The manipulator control is manual and visually guided; the sensor system supplements the visual information for control. The sensor system has been tested at the JSC Manipulator Development Facility under realistic conditions for grasping static and for capturing moving targets. More than 110 test runs have been performed by four operators, including an astronaut. The operational ground tests were very successful. The paper summarizes the features of the tests and evaluates the test results. The sensor system, together with its hardware and software components are also described. Future development and test plans are briefly outlined, including the use of an extended version of the sensor/display system.

I. INTRODUCTION

In visually guided manipulator control the operator always faces a fundamental problem: to assess the value of line-of-sight distance errors between end effector and object. In many cases the assessment of line-of-sight distance errors has an effect on the control of two or three task variables. Dependent on the viewing angle of the operator and on the construction of the end effector, the assessment of line-of-sight distance errors can influence the control of different combinations of task variables. A typical combination is the control of range, pitch and yaw errors of the end effector relative to the target object as indicated in Fig. 1. This combination of control errors prevails when the natural grasp plane of the end effector is perpendicular or near-perpendicular to the line-of-sight.

The visually guided control of the space shuttle Remote Manipulator System (RMS) end effector falls into the problem category indicated in Fig. 1. Actually, Fig. 1 shows the main contour of a four-claw end effector developed at the NASA Johnson Space Center (JSC) for a 15.2-m (50-ft) long manipulator which simulates the functions of the shuttle RMS.

A proximity sensor system has been developed at the Jet Propulsion Laboratory (JPL) and integrated with the four-claw end effector of JSC. The sensor system consists of four proximity sensors. The sensor system was designed to supplement the visual information for control in providing guidance data to the operator near the grasp envelope of the end effector where visual perception of depth, pitch and yaw errors are poor. The use of the sensor system is presently restricted to the verification of a "successful grasp state" of the end effector before the operator initiates grasp action. The "successful grasp state" is defined by the dimensions of the end effector's grasp envelope and by the dynamics of grasp. Geometrically, the "successful grasp state" is any allowable combination of depth, pitch and yaw errors which guarantees the mating of all four claws with the grapple fixture. When a "successful grasp state" has been reached, the sensor data processing electronics automatically turn on a simple "success display" (a buzzer or a green light, or both), indicating to the operator that he is ready to grasp.

Very successful operational ground tests have been conducted with the sensor and simple display system at JSC under realistic payload handling conditions to grasp standing and to capture moving targets. Altogether 112 test runs have been performed by four operators. The final result is that, when the "success display" was on, the operators achieved a capture every time. There were no operator mistakes under sensor-indicated grasping conditions, and the sensors never indicated wrong conditions for grasping.

The sensor system is described in Section II. The hardware and software details of the sensor system can be found in Appendix A and B. The test scenarios, including a brief description of the JSC Manipulator Development Facility (MDF) test equipment, are described in Section III.

The test data and their evaluation are presented in Section IV. Future plans are outlined in Section V, and the conclusions are summarized in Section VI.

II. PROXIMITY SENSOR SYSTEM

1. Concept.

The top surface of the grapple fixture defines the grasp plane of the target object. The four outer hinges of the four-claw end effector define the grasp plane of the end effector. Thus, the stated terminal guidance and control problem is equivalent to the control of depth positioning and pitch and yaw alignments of the end effector's grasp plane relative to the grapple fixture's grasp plane. But depth positioning and pitch and yaw alignments of two planes are in a natural way related to range measurements between the two planes along the normal of one of the two planes taken at several (at least three) points of that plane. Of course, all the measurement points cannot be located along the same line on the plane.

Since it is an advantage to utilize the square symmetry of the four-claw end effector for the given guidance/control problem, a square matrix of four electro-optical proximity sensors has been developed. (Electro-optical proximity sensing has been described in Ref. 1.) The sensors are mounted to the center square frame of the end effector which is not part of the closure mechanism. The sensors are mounted so that the beam (or optical axis) of each sensor is parallel to the roll axis of the end effector which is at the center of the square frame. Using the claws to define the corners of a square, the sensors are mounted on the midpoints of the sides of the square. (See Fig. 2.) In this way, the two lines connecting sensors on opposite sides of the square frame intersect orthogonally at the center of the square, i.e., on the roll axis of the end effector. Hence, the two orthogonally and centrally intersecting lines define a natural frame for sensing pitch, yaw and depth errors from the end effector relative to the top plane of the grapple fixture. (See Fig. 3.)

This emplacement of the sensors assures a permanently open, unobstructed field of view for the sensor beams relative to the grapple fixture's top surface and does not interfere with the operator's visual field. At the same time, this sensor emplacement also protects the sensors from incidental collisions with the grapple fixture.

The total sensitive range of each sensor is about 18 cm. (7 in.), thus covering somewhat longer range than the depth of the grasp envelope which is about 10 cm. (4 in.). The sensor resolution in range measurement is about 2.5 mm. (0.1 in.). The separation between two sensors on the same axis is about 14 cm. (5.5 in.). Hence, the angular measurement resolution is about 1 deg. More on the sensor can be found in Appendix A.

The application of these sensors to the described guidance/control problem requires that the top of the grapple fixture have a continuous flat surface with known and homogeneous reflectivity.

2. Implementation.

The major components of the sensor system and their configuration relative to the JSC MDF are indicated in Fig. 4. Some details of the electronic instrumentation are described in Appendix A. The complete system (without the microprocessor unit) is shown in Fig. 5. The microprocessor together with the "cabin electronics" is shown in Fig. 6.

The data processing required to drive the "success display" has two modes: analog and digital. In both cases the "success display" is a buzzer or a green light. When the buzzer and/or green light is on, it indicates to the operator that depth position and pitch and yaw alignments of the end effector relative to the grapple fixture are within allowable limits for successful grasp. Thus, the "success display" does not indicate to the operator the details of the three-dimensional (depth, pitch and yaw) error states. It only indicates the allowable combination of these three independent errors for successful grasp. Assuming 0.1 in. resolution for depth sensing and 1 deg. resolution for pitch and yaw alignments sensing, and assuming 3 in. total allowable depth error and 10 deg. total allowable error for both pitch and yaw alignments individually, we then have 30 by 10 by 10, that is, 3000 allowable combinations of depth, pitch and yaw errors for successful grasp. The simple "success display", therefore, lessens the operator's perceptive and cognitive workload by relieving him from the real-time task of evaluating the details of the depth, pitch and yaw error states. The drive logic of the "success display" does this evaluation for the operator automatically. The operator receives only a simple binary information from the display: yes or no.

The analog drive logic implementation is quite simple. (See Appendix A.) In fact, in this simple analog implementation the full capabilities of the sensor system cannot be utilized to account for all physically possible combinations of depth, pitch and yaw error states which, due to the dimensions of the end effector's grasp envelope, still would allow a successful grasp. The primary reason for the limitation of the simple analog logic is the nonlinearity of the sensor response.

To achieve a full utilization of the sensor system capabilities, a digital unit has been employed for sensor data processing and for driving the simple "success display". The digital unit is an Intel 80/20-4 single-board microcomputer with an Intel single-board A/D converter.

For the purpose of experimentation, several "success algorithms" have been implemented in the microcomputer to drive the on-off "success displays". The algorithms are simple and account for all (or for almost all) allowable error state combinations for successful grasp. Algorithms have also been implemented which only utilize the signals of any three out of the four sensors to indicate the allowable error combinations for successful grasp. This is useful if one sensor eventually fails or if one sensor beam eventually misses the top (reference) surface of the grapple fixture due to the allowable lateral alignment errors.

It is noted that the four-sensor configuration shown in Fig. 2 is really redundant to define and compute depth, pitch and yaw errors.

A triangular configuration of three sensors would be sufficient for that purpose. The redundancy of the four-sensor system is obvious from the fact that depth error can be computed in two independent ways as shown in the lower right part of Fig. 3. Hence, if one of the four sensor signals is missing, it can be reconstructed from depth error data and, consequently, both pitch and yaw errors can be computed as if all four sensor signals were present.

Three kinds of basic "success definitions" have been developed, each with three sets of "success parameters". All nine variations have been implemented for "all four" and for "three-out-of-four" sensors. Altogether 18 algorithms are stored in Erasable, Programmable Read-Only Memory (EPROM) in the microcomputer. Any one of the 18 algorithms is easily callable by dialing the appropriate number between 1 and 18 on a Binary Coded Decimal (BCD) switch integrated with the microcomputer. Basic algorithm no. 1 is called "simple prism", no. 2 is called "prism extended by pyramids", and no. 3 is called "cone". More on the algorithms and on the sensor/display software can be found in Appendix B.

III. GROUND TESTS

Testing of the proximity sensor augmented end effector was performed during May 1978 in the Manipulator Development Facility (MDF) at the NASA Johnson Space Center (JSC).

1. Test Facility and Equipment

The MDF (shown in Fig. 7) utilizes a full-scale mockup of the Shuttle Orbiter aft crew cabin and payload bay, a Remote Manipulator System (RMS) workstation, and arm assembly, thereby offering unique testing capability, particularly in relation to the RMS visual system which includes the closed circuit television (CCTV) and direct viewing by the operator. Primary functions of the MDF include serving as an engineering design tool for the RMS, assessing the man/systems interfaces, developing operating procedures for the RMS activities, and crew training. Minor differences in design and operational capabilities are necessary in order to operate the MDF manipulator system in a one-g environment rather than in zero-g. The MDF also features a large air bearing floor [17 by 24 m. (56 by 80 ft.)] which supported the dynamic sensor testing with an air bearing payload simulator. (See Fig. 8).

The manipulator arm, 15.2 m. (50 ft.) long and hydraulically actuated, can be controlled in six degrees of freedom with resolved rate similar to the flight RMS. The arm, originally delivered as a commercial item from General Electric, can be operated via a position-control system through a replica arm or via resolved rate command using two hand controllers. Replica arm commands are transmitted as individual joint voltages to a comparator network which outputs command-minus-feedback error voltages as drive signals to the joint servos. The replica system is presently maintained as a back-up to the rate system.

The resolved rate command system calculates joint rates required to satisfy the end effector rates commanded by two three-degree-of-freedom proportional hand controllers. These joint rates are numerically integrated and added to the feedback joint angles to generate commanded joint angles. The corresponding joint voltage is sent to the same comparator network utilized with the replica system. The result is a resolved rate achieved with a position-control hardware system. A Systems Engineering Laboratories (SEL) 32/35 computer with 256K bytes of 900 nanosecond memory is used in the rate system emulation. (See Fig. 9).

Various coordinate reference frames are simultaneously calculated in the SEL for use in different applications when selected by the operator. The end-effector-referenced and Orbiter-referenced frames (Figs. 10 and 11) were used during this evaluation--the former for tasks in which only CCTV was utilized (i.e., no direct vision) and the latter for tasks primarily involving direct vision. The end effector system is basically camera-referenced: a controller input (e.g., translation hand controller "left" command) causes end-effector movement from right to left on the TV monitor. Similarly, a translation hand controller command to starboard in the Orbiter frame causes end-effector movement from port to starboard along a line through the end effector tip and parallel to the Orbiter Y axis.

2. Test Scenario

The test was divided into two phases: static runs which were completed with a non-moving payload, and dynamic runs performed with a moving payload on the air bearing floor of the MDF. In the static set of runs, the operator was located at the RMS workstation in the cabin area. The payload with special grapple fixture was located in the payload bay in one of two attitudes: grapple fixture on top and front. Given a go-ahead, the operator maneuvered the end effector using direct vision and/or closed circuit television feedback to effect an alignment with the grapple fixture. The sensor utility was limited in this case since accuracy was the critical factor to the operator.

However, during the dynamic phase, the sensor became very useful as the operator was primarily concerned with payload capture rather than alignment accuracies. In these runs, the operator was located both in the cabin--in which case he used television feedback to make the capture--and on the air bearing floor--in which case he used direct vision to make the capture. In both cases, the sensor was used to provide initial indication that within-tolerance grapple conditions had been achieved.

Figure 12 shows the sensor-augmented end effector on the Shuttle mock-up manipulator and a few test scenes at the JSC MDF.

IV. TEST DATA AND DISCUSSION

Four subjects completed 20 static and 92 dynamic runs. In the static configuration, operators were 100% successful at achieving end effector alignment with the grapple fixture within the grapple envelope. Mean and maximum linear and angular misalignments for these runs are presented in Table 1. Subjective evaluations indicated that the sensor's utility was not fully exercised during this test since the goal was accurate alignment, not just achieving valid grapple conditions (i.e., on the fringes of the envelope). However, operator comments indicated that if the sensor feedback included quantitative accuracy information, the sensor potentially could allow a very accurate alignment, and, thus become a useful tool in achieving a minimum-load grapple.

For the dynamic case, the operators were 100% successful in payload capture when the sensor feedback was positive (i.e., valid grapple conditions achieved) as indicated by sensor display and aural tone. Mean and maximum linear and angular misalignments for these runs are presented in Table 2. Mean times to grapple as a function of payload speed are presented in Table 3. Grapple success data are shown in Table 4. While mean linear and angular misalignments were similar for both static and dynamic cases, the variance was higher for the dynamic case. The reader will note that while the sensor was 100% reliable, grapple success was more readily achievable when the operator had direct visual access to the payload as opposed to viewing the payload via the television system. Subjective comments on the sensor for the dynamic runs indicated that the sensor was extremely useful for the payload capture case. The sensor also seemingly allowed the operator to grapple before the other present visual cues (e.g., target) indicated that allowable alignment conditions had been met. Generally, the subjects felt that the test demonstrated the desirability and utility of the proximity sensor feedback as an aid to the RMS operator during payload grapple activities.

In general, subjects felt more "comfortable" with the sensor when compared to a target which presented visual cues only. With the proximity sensor, it is possible that some visual workload was transferred to the audio channel thereby distributing the workload and reducing tension in the subjects to some degree. Additionally, the operators could wait for either an audio or visual "discrete" signal indicating within-tolerance grasp conditions, which is clearly a much easier task than qualitative and continuous assessment of a passive visual target.

V. FUTURE PLANS

As pointed out above, the simple "success display" does not show the details of the three-dimensional (depth, pitch and yaw) error states. Advanced graphics display concepts have subsequently been implemented to convey to the operator not only the "success" information but also

the details of the three (depth, pitch and yaw) errors so that the operator will know from the sensors "what to do" in order to get the end effector to the "successful grasp state" or to fine-control the grasp. Figure 13 shows an advanced concept implemented in color graphics. Here, "success" is indicated by all error bars turning green. The unsuccessful error combinations are indicated by all error bars turning red. The length of the error bars is proportional to the respective errors under both "green" and "red" conditions. A new display based on the proportional bar technique and event indicator will be tested at the JSC MDF later this year.

Plans also include extending the sensor system with a longer range sensing component to provide guidance information to the operator ahead of the grasp envelope. It is also feasible to build a sensor system which would provide terminal guidance information in all six dimensions, that is, in three position and in three alignment degrees of freedom. The plans also include the potential redesign of the electronic implementation by incorporating fiber optic cables into the sensing system. This would greatly simplify flight implementation and would considerably improve signal quality.

VI. CONCLUSIONS

The tests have shown that proximity sensor information is a very effective and desirable aid to capture moving targets. It also has the potential to minimize grapple pre-load in grasping large or any static targets. The operators' subjective comments confirmed the usefulness and practicality of the proximity sensor based terminal guidance system concept supplementing the visual information for the control of grasp/capture operations. The tests have also indicated that extending the range of proximity sensing and providing more detailed but coordinated display information to the operator would further enhance the utility of the sensor system.

ACKNOWLEDGMENT

The contribution of the following persons at JPL to the work described in this paper is gratefully acknowledged: P. L. Cassell (sensor electronics), W. A. Hermann (sensor optics), G. Paine (digital system and graphics display), and E. Shalom (software). The supporting interest of Division Chief A. J. Louviere of JSC is also gratefully acknowledged.

The research described in this paper was partly carried out at the Jet Propulsion Laboratory, California Institute of Technology, under NASA Contract NAS7-100.

APPENDIX A.

SENSOR HARDWARE

The optical configuration of the proximity sensor head is shown in Figure 14 together with a few design data. 80-mm focal length collimating lenses are used in the sensor head. The characteristics of sensor output and sensor head location relative to the end effector's grasp range are shown in Figure 15. Note that the top of the bell-shaped signal curve is located inside the unreachable space. In this arrangement, only the outer leg of the signal curve can be utilized for distance sensing. Consequently, the sensed distance is always a single-valued function of the voltage output of the sensor. The signal curve shown in Figure 15 is the actual calibration curve of one of the four sensors relative to a dull reddish reflecting surface.

The main elements of the electronic instrumentation are shown in Figure 16. Block "B" in Figure 16 (the display drive logic) can be an analog or a digital unit. The analog drive logic implementation of the success display is explained in Figure 17.

APPENDIX B.

SENSOR AND DISPLAY SOFTWARE

The software implemented in assembly language in an Intel 80/20-4 single-board microprocessor performs three major functions in real time:

(1) Linearization of sensor data.

(2) Determination of range, pitch and yaw errors and their combinations relative to "successful grasp criteria."

(3) It drives the display.

Sensor data linearization is accomplished by table look-up. The calibration curve for each sensor is stored in memory for 0 to 4.2 in. (0 to 11 cm.) range with 0.05 in. (about 1.25 mm.) separation between data points. Hence, each sensor calibration curve is represented by 84 data points.

For the purpose of experimentation, three different algorithms have been defined and implemented for "successful grasp." The computer program organization is shown in Figure 18. The three "success algorithms" are explained in Figure 19, together with the appropriate symbol definitions.

The difference between the three algorithms is as follows. Algorithm No. 1, called "simple prism algorithm", defines the successful grasp volume as a prism. It looks at the four sensor ranges individually.

Each sensor range must be between predefined upper and lower limits corresponding to the allowable depth error. The limits are identical for all four sensor ranges. At the same time, the differences between the two sensor ranges on both the pitch and yaw axes must be less than a predefined value corresponding to the allowable pitch and yaw errors. (See also Figure 3.) Algorithm No. 2, called "prism extended by pyramids", defines the successful grasp volume as a prism extended with pyramids at the top and at the bottom of the prism. First, it computes the range. The range must be within predefined limits. Then, the allowable pitch and yaw errors are computed and gated as functions of the range according to the trapezoid formula shown in Figure 19. Algorithm No. 3, called "conic algorithm", defines the successful grasp volume as a cylinder extended with cones at the top and at the bottom of the cylinder. First, it computes the range like Algorithm No. 2. But, it combines the individually allowable pitch and yaw errors into a single allowable cone error condition which is gated as a function of the range according to the trapezoid formula shown. Algorithm No. 3 is the most powerful among the three algorithms. Figure 20 illustrates the successful grasp volumes as defined by the three different algorithms.

It is noted that the redundancy of the four-sensor system versus depth, pitch and yaw errors is obvious from the fact that depth error can be computed in two independent ways as shown in Figure 3. Hence, if one of the four sensor signals is missing it can be reconstructed from range data and, consequently, both pitch and yaw errors can be computed as if all four sensor signals were present.

Each "success algorithm" has been implemented with three sets of "success parameters". Further, all nine variations have been implemented for "all four" and for "three-out-of-four" sensor cases. Altogether 18 algorithms are stored in the microcomputer. The parameters with the corresponding switch number assignments are listed in Table 5. The parameter symbols of Table 5 follow the definitions explained in Figure 19.

The computer program subroutines shown in Figure 18 perform the following functions. The first subroutine initializes the Intel 8255 parallel peripheral interface ports. The second subroutine initiates the A/D conversion process, reads and stores the output voltage data of each of the four sensors. The third subroutine identifies the ranges sensed by each of the four sensors by searching through the calibration table. The fourth subroutine looks for missing distance values. The fifth subroutine is the central routine for executing the algorithmic computations. The last subroutine examines the result of the algorithmic computations and sets the light and buzzer (the "success display") accordingly. The execution of the whole program takes only a few milliseconds.

Table 1. Static misalignments

		MISALIGNMENT	
		MEAN	MAXIMUM
POSITION:	X	1.1 IN.	1.7 IN.
	Y	0.4 IN.	0.8 IN.
	Z	0.3 IN.	0.9 IN.
ATTITUDE:	PITCH	2.2^0	2.2^0
	YAW	1.7^0	3.5^0
	ROLL	2.1^0	5.5^0

Table 2. Dynamic misalignments

		MISALIGNMENT	
		MEAN	MAXIMUM
POSITION:	X	0.6 IN.	1.6 IN.
	Y	1.0 IN.	2.7 IN.
	Z	0.7 IN.	4.0 IN.
ATTITUDE:	PITCH	2.9^0	6.8^0
	YAW	2.2^0	6.0^0
	ROLL	3.2^0	9.5^0

Table 3. Dynamic grapple times

	VISUAL ACCESS	
PAYLOAD SPEED (FT/SEC)	TV ONLY	DIRECT VISION ONLY
0.1	1.35 MIN.	0.95 MIN.
0.2	1.23 MIN.	0.92 MIN.
0.3	1.30 MIN.	0.72 MIN.

Table 4. Dynamic grapple success

PAYLOAD SPEED (FT/SEC.)	VISUAL ACCESS — TV ONLY	VISUAL ACCESS — DIRECT VISION ONLY
0.1	100%	100%
0.2	88%	100%
0.3	35%	100%

Table 5. Parameter values and panel switch assignments of "success algorithms"

PANEL SWITCH ASSIGNMENTS

SWITCHES	ALGORITHM SELECTED	
01	SIMPLE - PARAMETER SET 1 (SM1)	FOUR SENSORS
02	SIMPLE - PARAMETER SET 2 (SM2)	
03	SIMPLE - PARAMETER SET 3 (SM3)	
04	EXPANDED - PARAMETER SET 1 (EX1)	
05	EXPANDED - PARAMETER SET 2 (EX2)	
06	EXPANDED - PARAMETER SET 3 (EX3)	
07	CONIC - PARAMETER SET 1 (CO1)	
08	CONIC - PARAMETER SET 2 (CO2)	
09	CONIC - PARAMETER SET 3 (CO3)	
10	SIMPLE - PARAMETER SET 1 (SMTF1)	ANY THREE OUT OF FOUR SENSORS
11	SIMPLE - PARAMETER SET 2 (SMTF2)	
12	SIMPLE - PARAMETER SET 3 (SMTF3)	
13	EXPANDED - PARAMETER SET 1 (EXTF1)	
14	EXPANDED - PARAMETER SET 2 (EXTF2)	
15	EXPANDED - PARAMETER SET 3 (EXTF3)	
16	CONIC - PARAMETER SET 1 (COTF1)	
17	CONIC - PARAMETER SET 2 (COTF2)	
18	CONIC - PARAMETER SET 3 (COTF3)	

PARAMETER SETS (Inches)

SWITCH NUMBERS		A	B	C
1 OR 10	SM1	0.8	2.8	1.0
2 OR 11	SM2	0.5	2.3	1.0
3 OR 12	SM3	0.1	2.8	1.0

		HA	HB	HC	HD	G	K
4 OR 13	EX1	0.5	1.0	2.3	2.8	1.1	2.2
5 OR 14	EX2	0.5	0.9	2.1	2.5	1.1	2.75
6 OR 15	EX3	0.1	0.5	2.4	2.8	1.1	2.75
7 OR 16	CO1	0.5	1.0	2.3	2.8	1.25	2.5
8 OR 17	CO2	0.5	0.9	2.1	2.5	1.25	3.125
9 OR 18	CO3	0.1	0.5	2.4	2.8	1.25	3.125

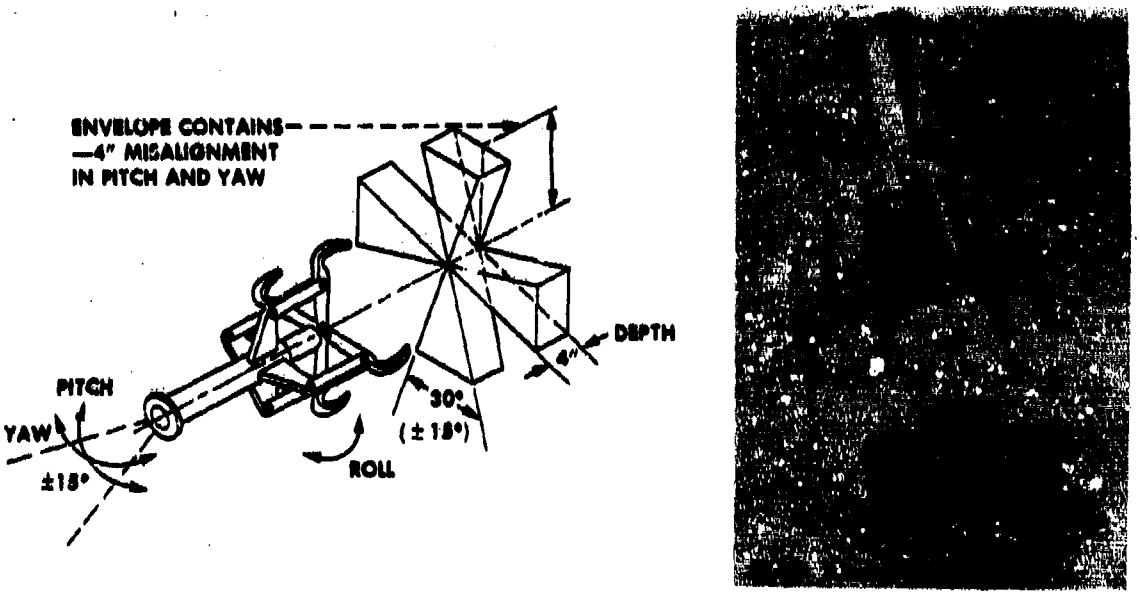


Figure 1. JSC four-claw end effector

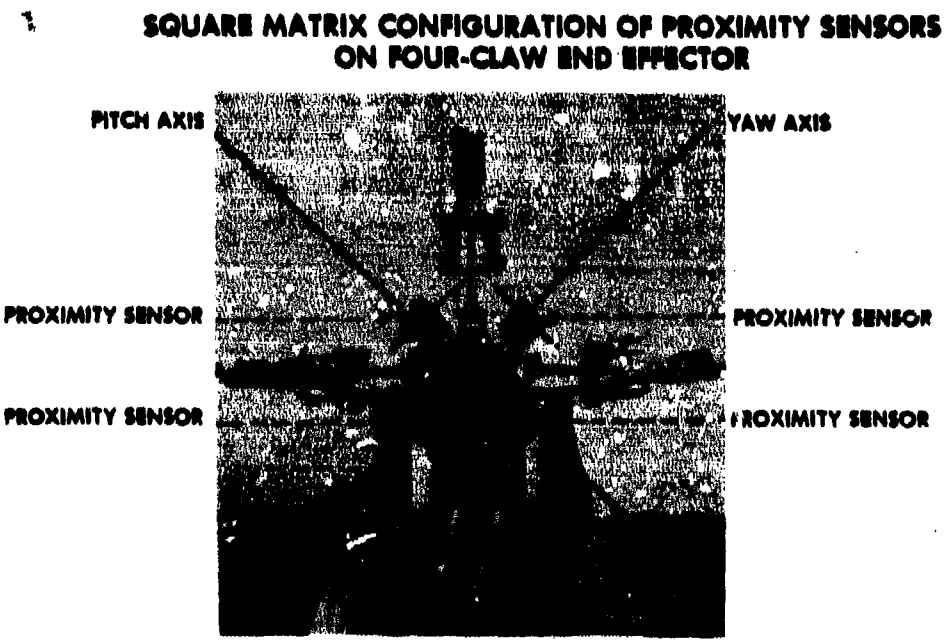


Figure 2. Four proximity sensors on four-claw end effector

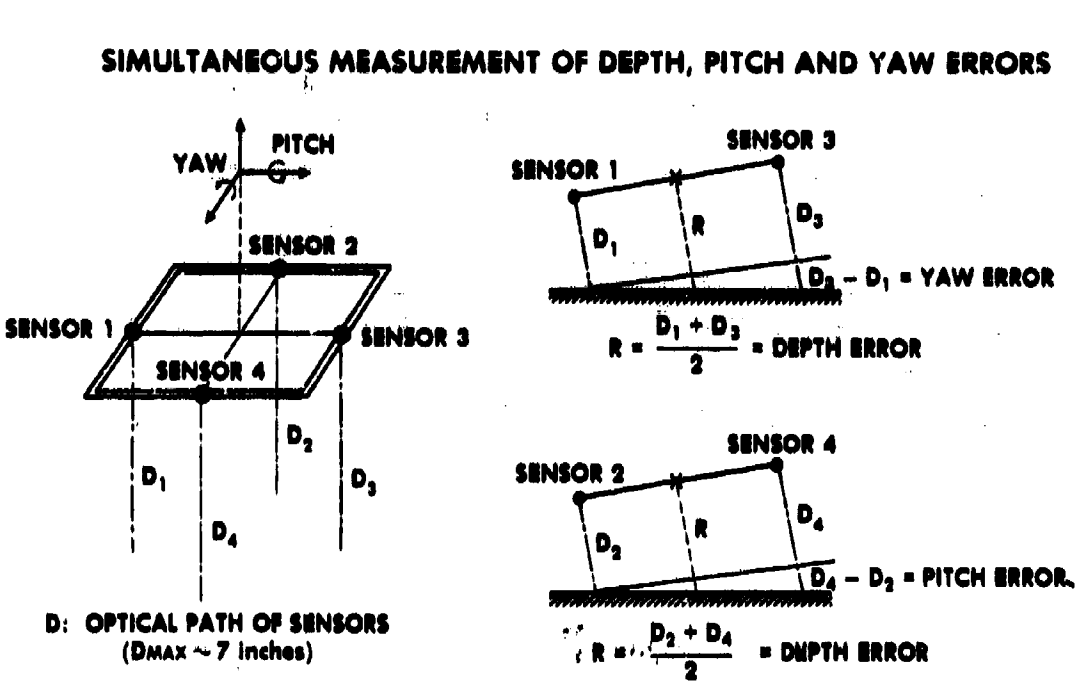


Figure 3. Four-sensor operation concept

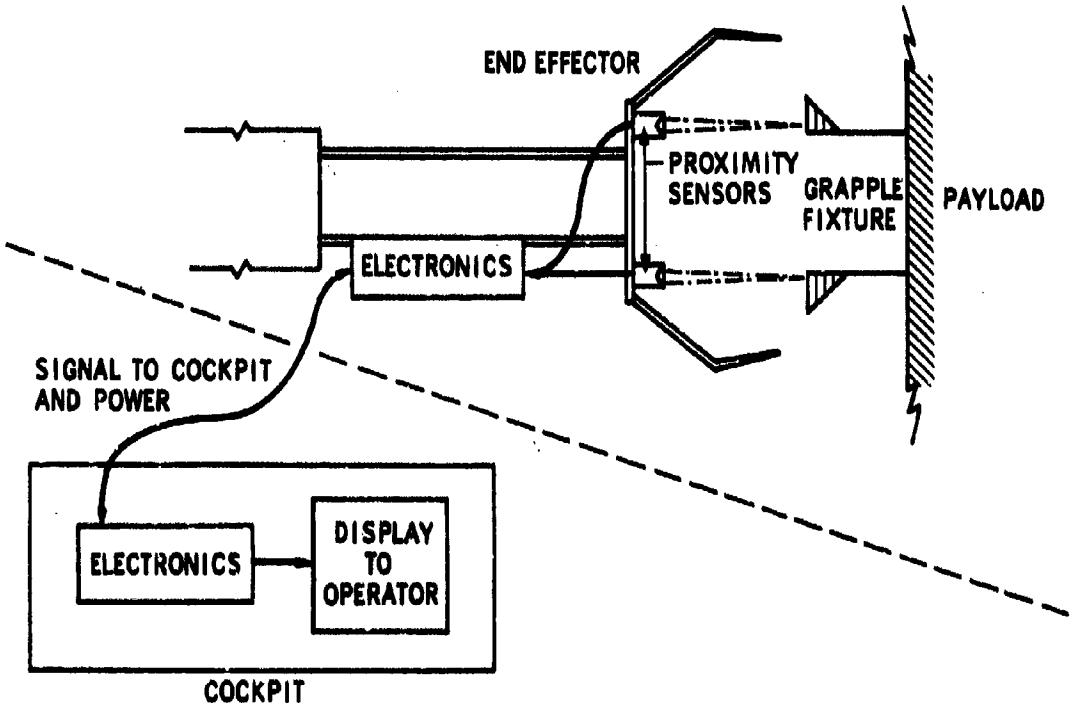


Figure 4. Sensor system components

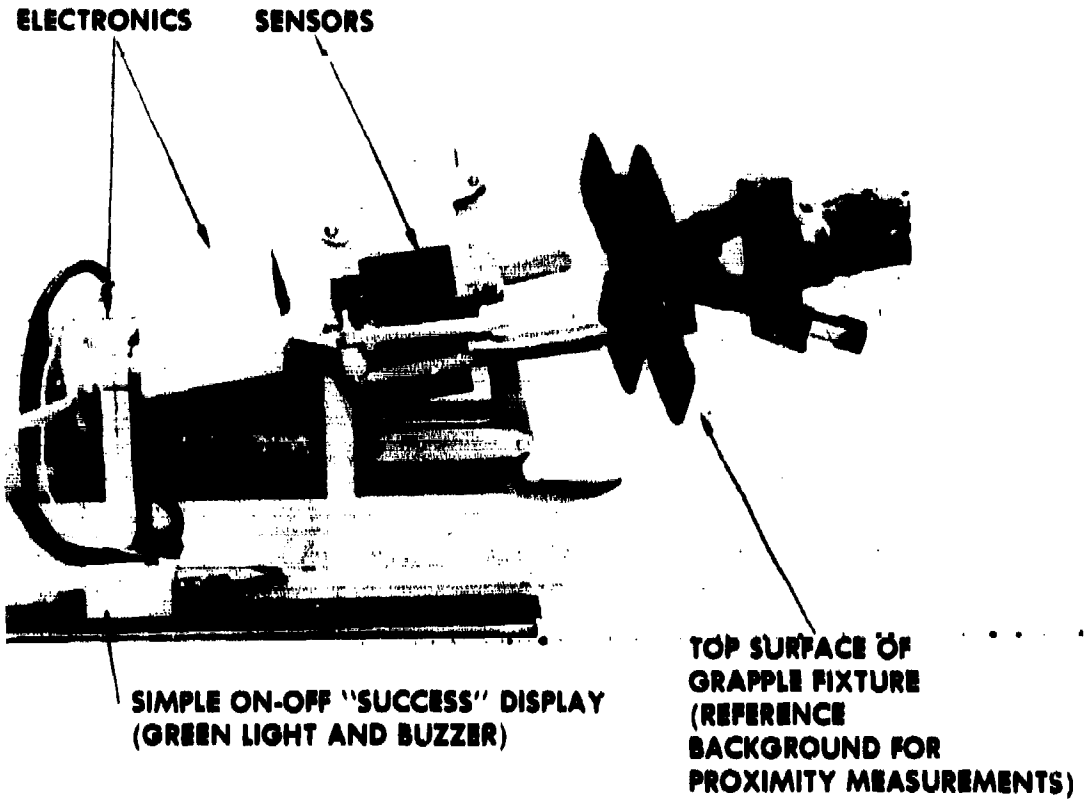


Figure 5. Overall proximity sensor system

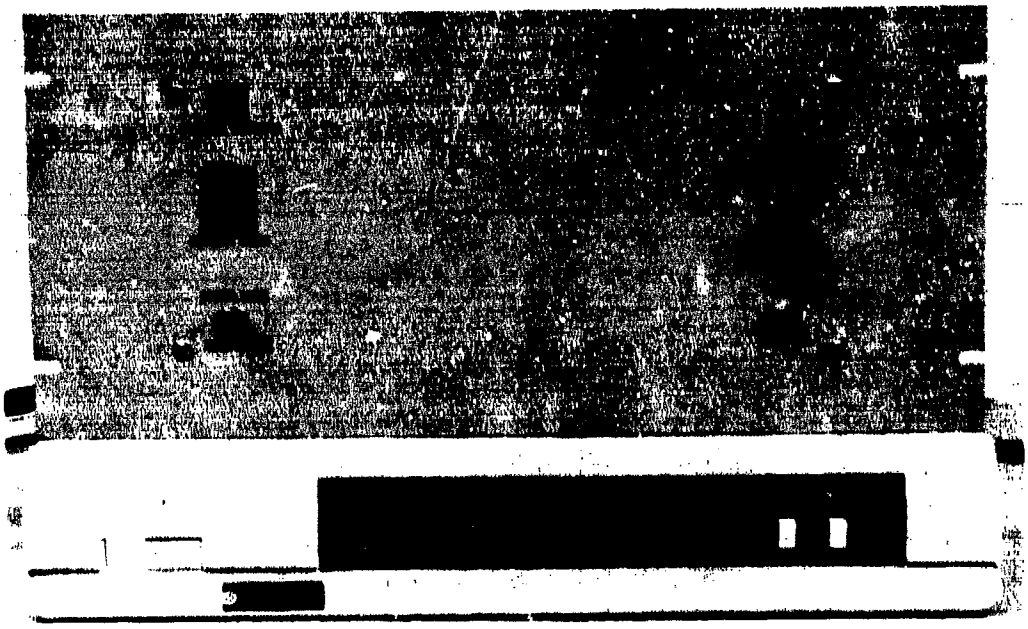


Figure 6. Microprocessor and drive electronics

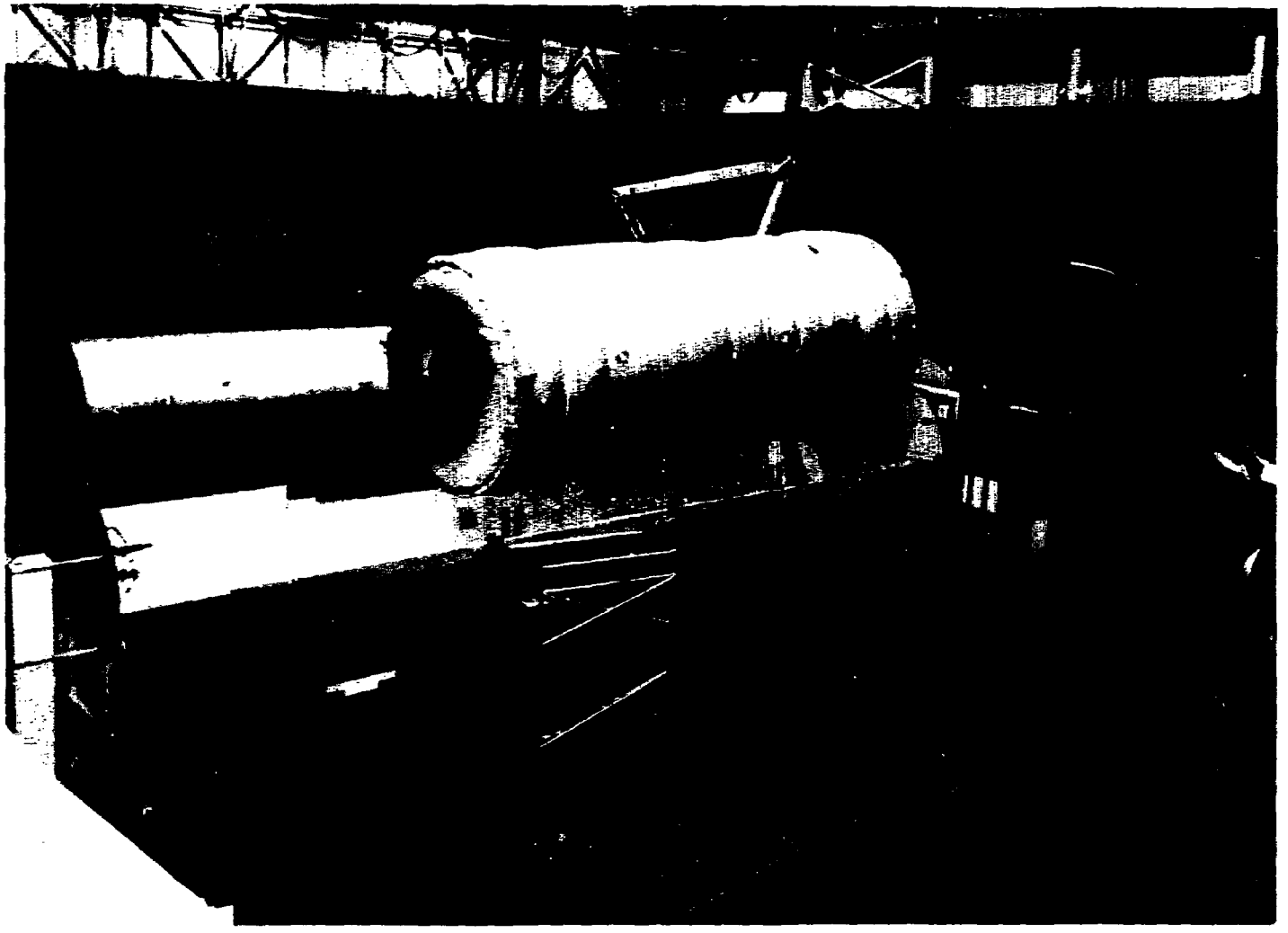


Figure 7. Shuttle mock-up at JSC MDF

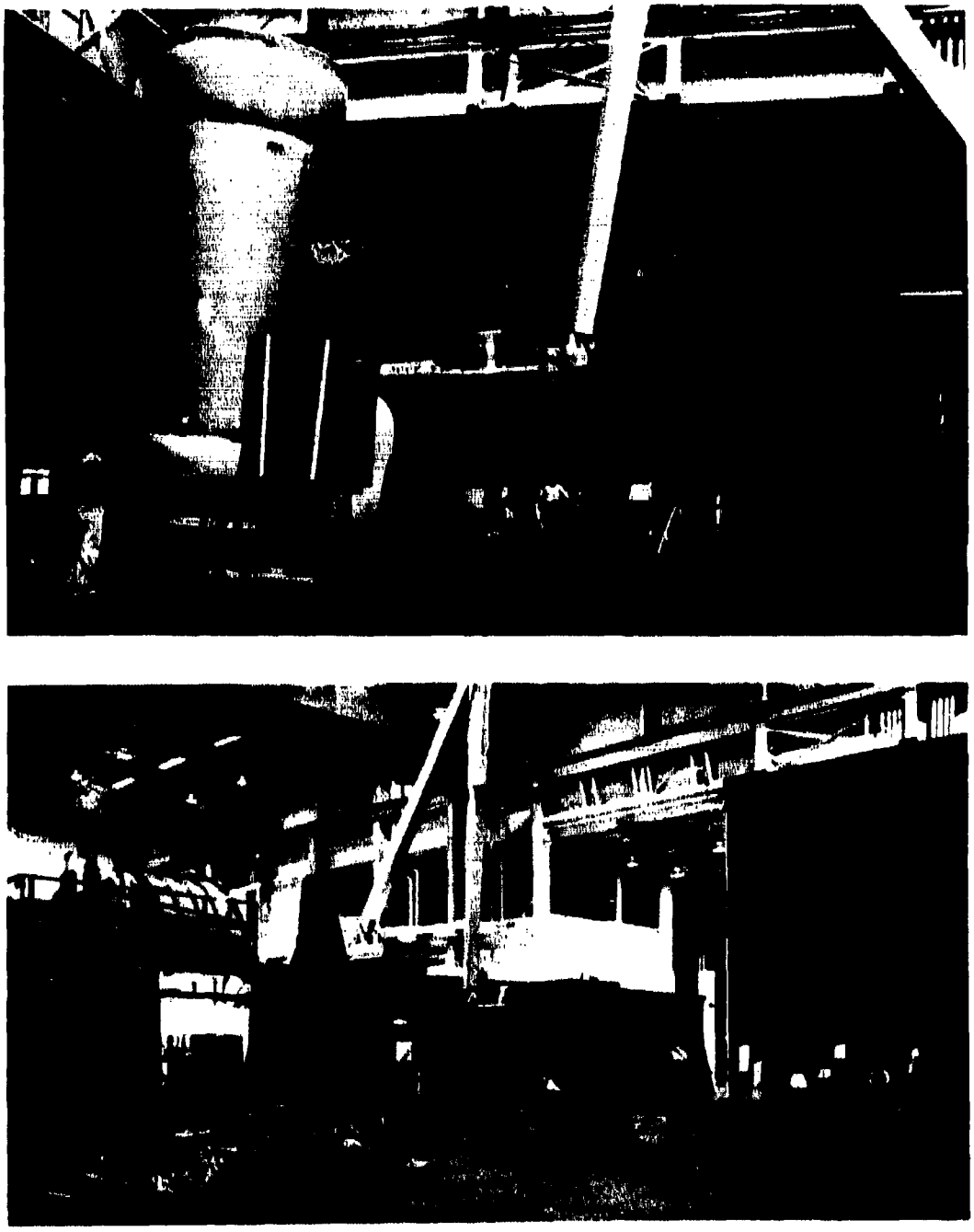


Figure 8. JSC MDF with air bearing floor

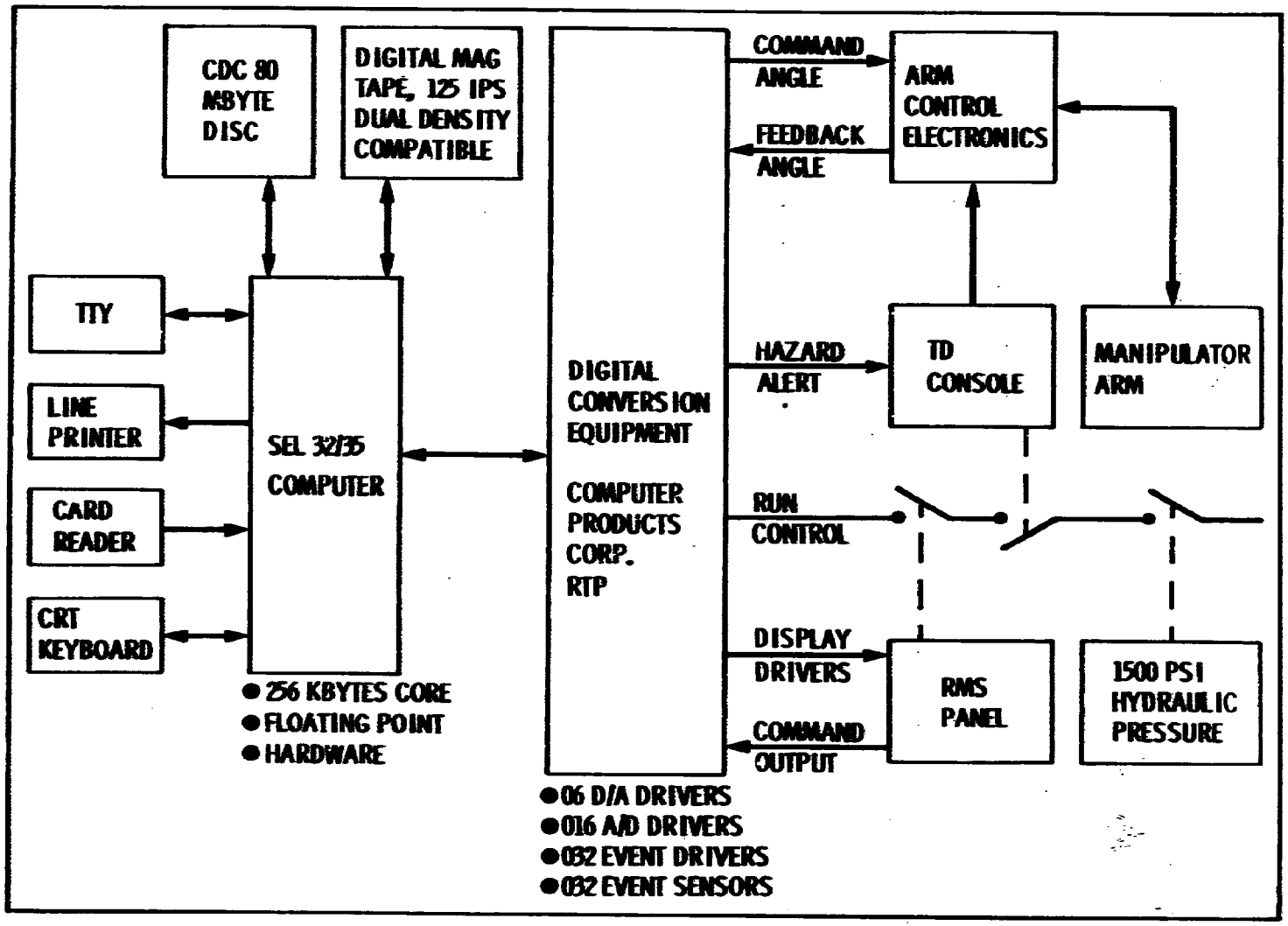


Figure 9. Control computer system at JSC MDF

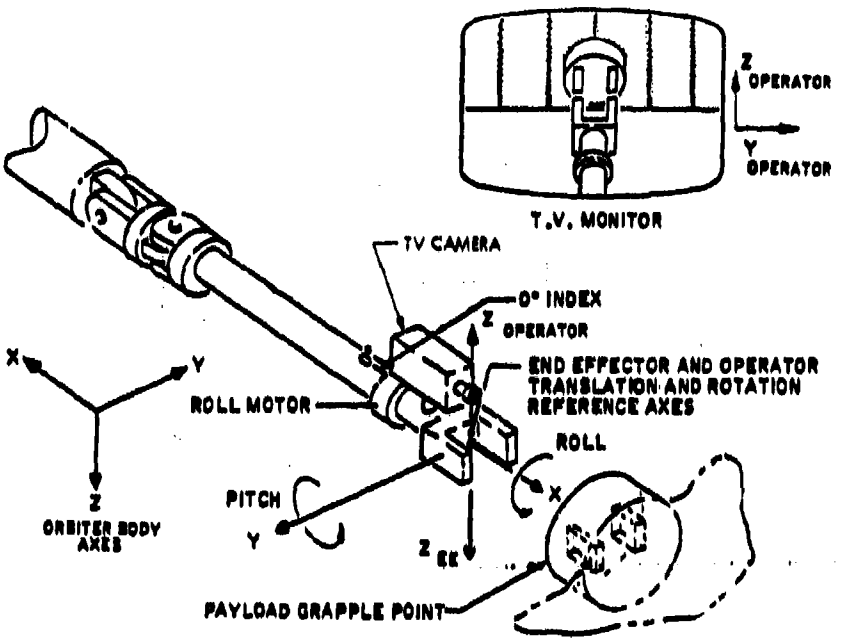


Figure 10. End effector reference frame for control

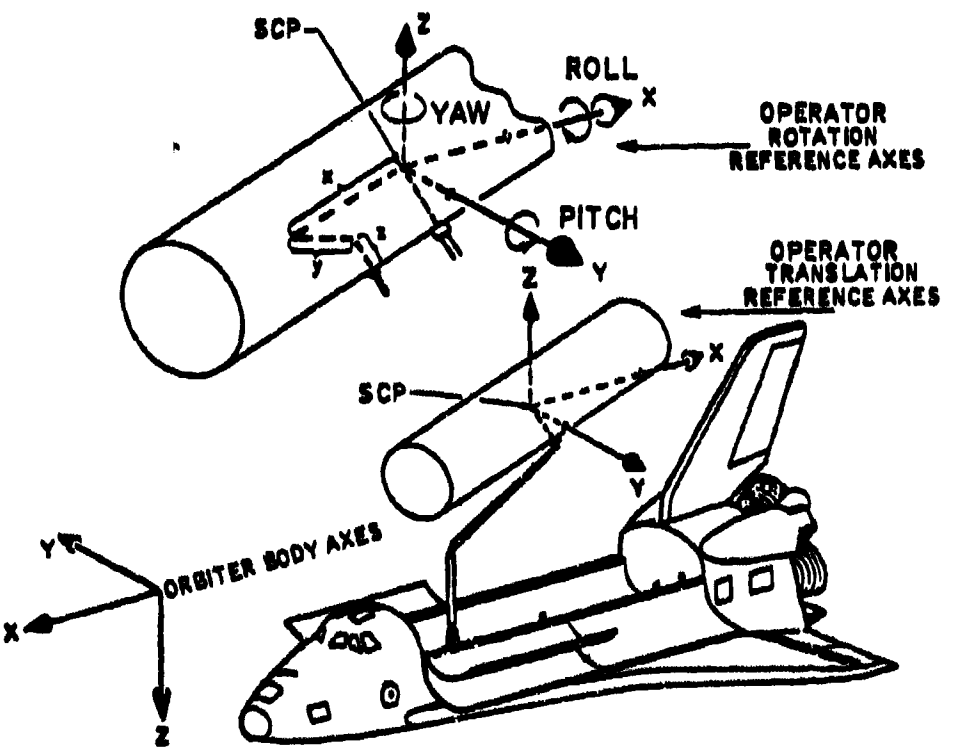


Figure 11. Orbiter reference frame for control

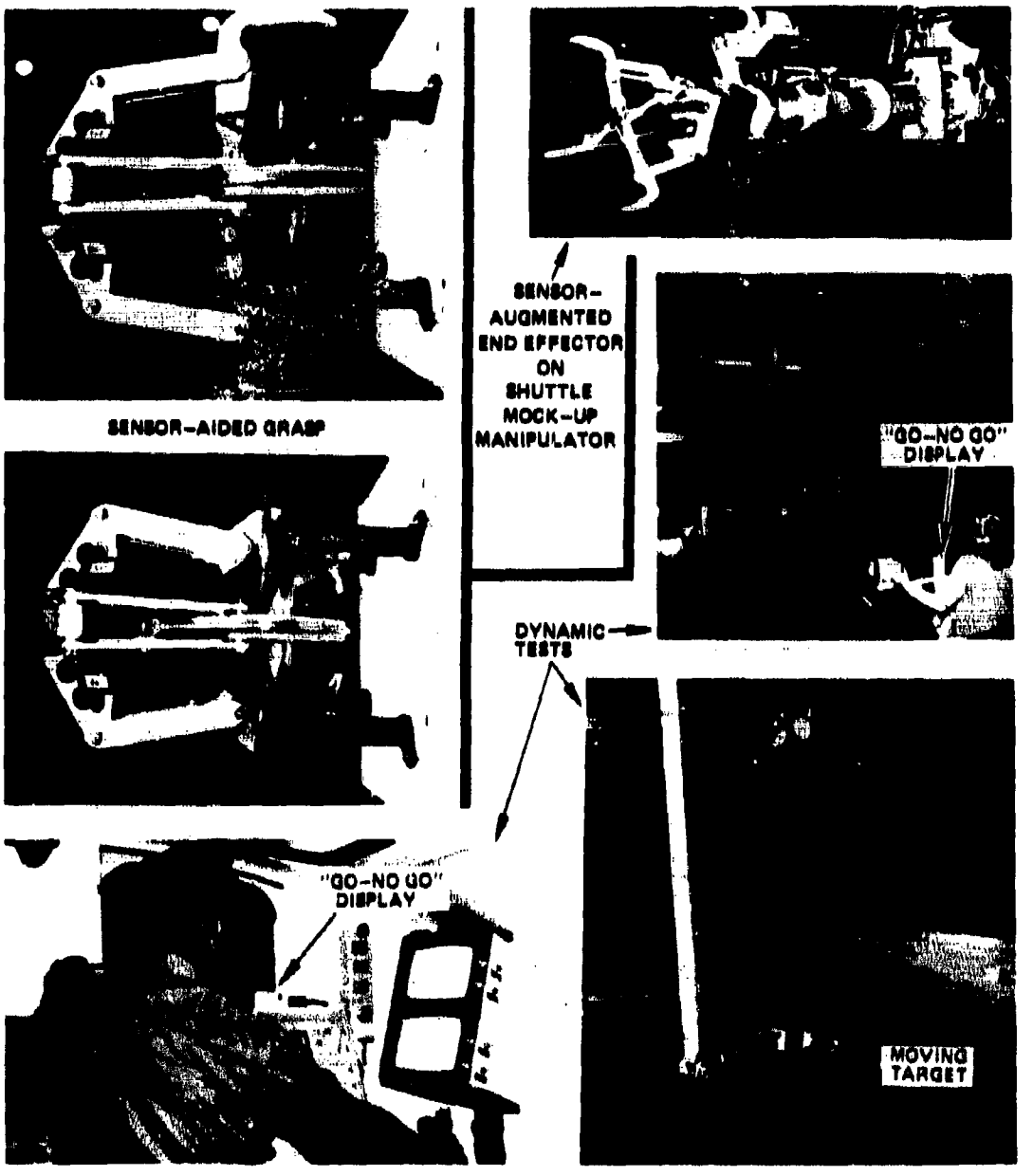


Figure 12. Proximity sensor test scenes at JSC MDF

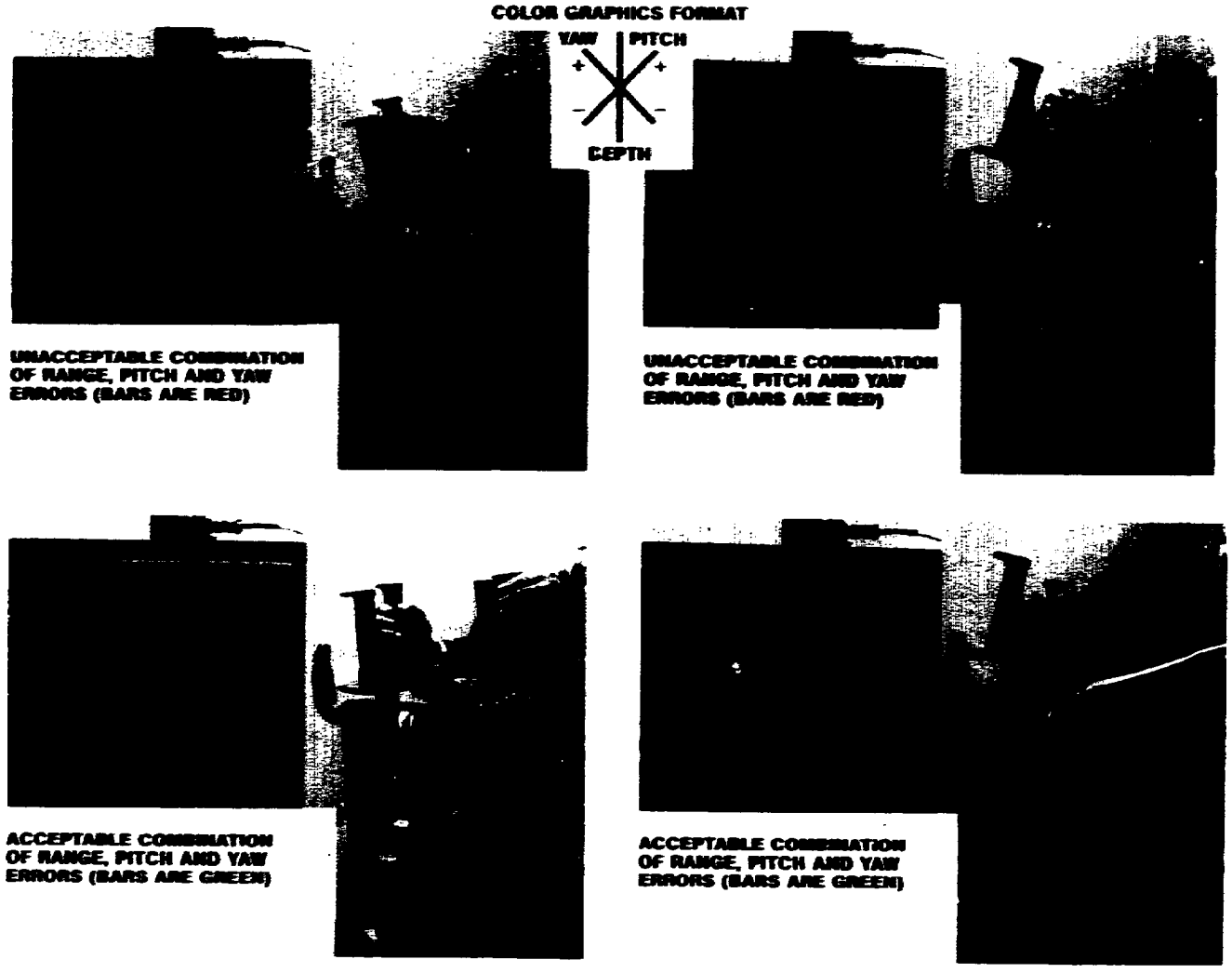


Figure 13. An advanced sensor display concept

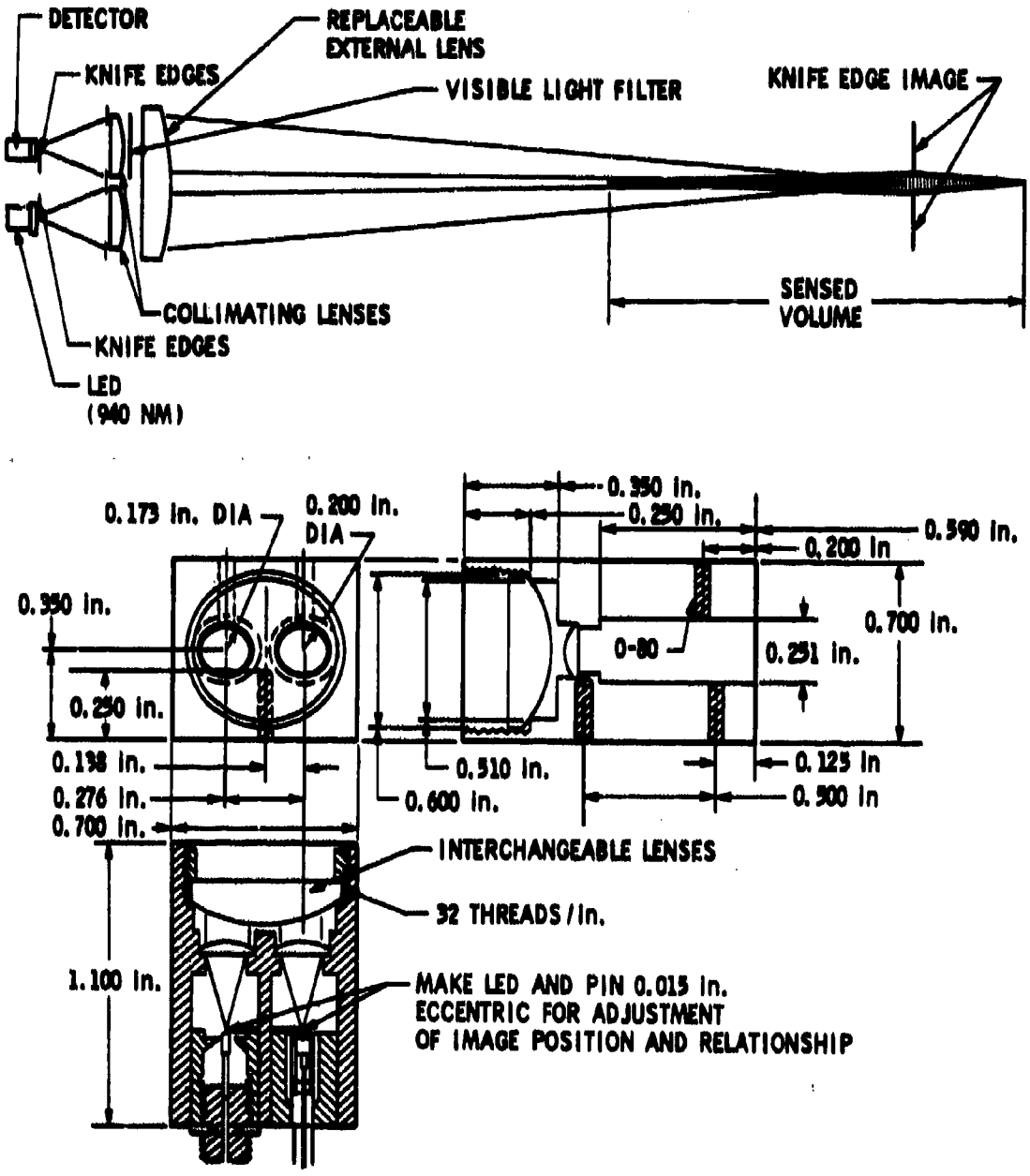


Figure 14. Optical configuration of proximity sensor

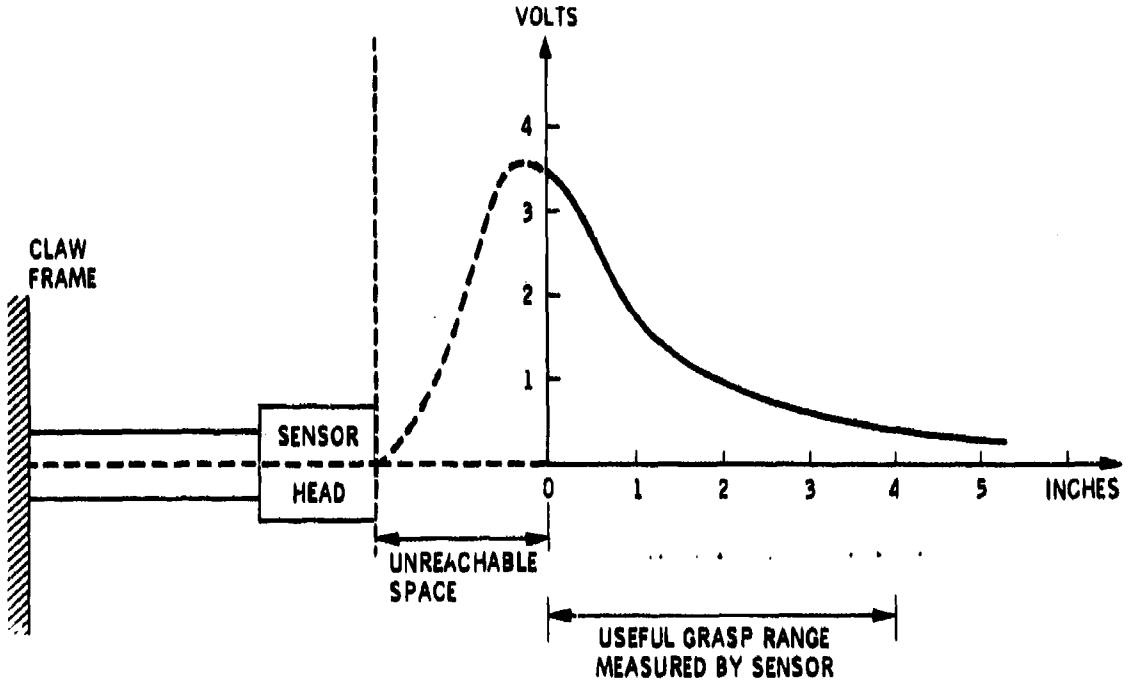


Figure 15. Emplacement and output signal of proximity sensor relative to grasp range

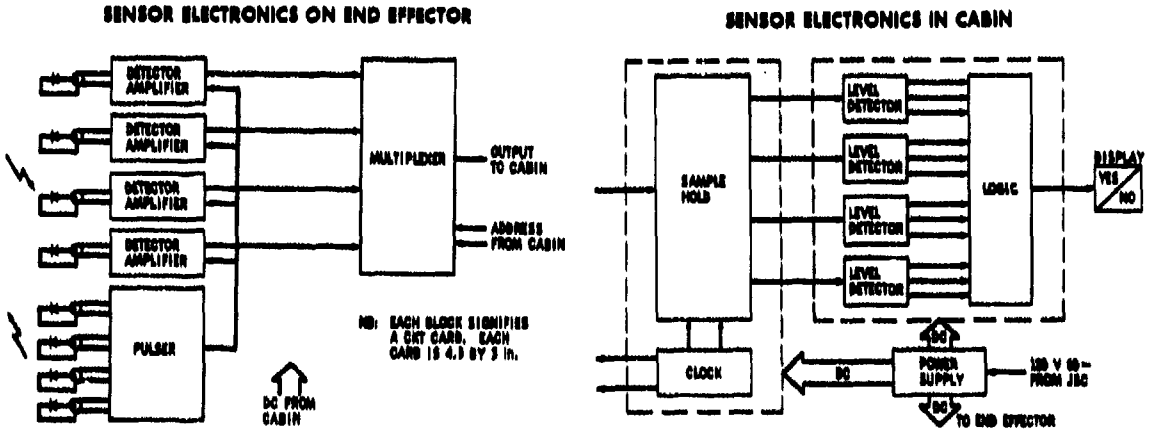
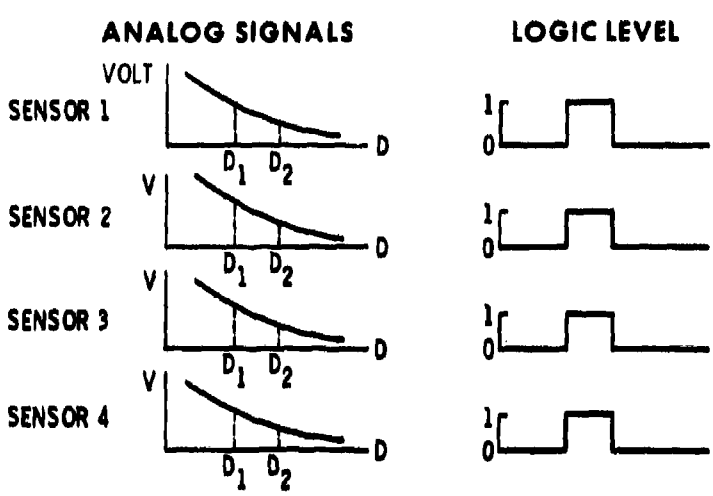


Figure 16. Proximity sensor electronics for JSC four-claw end effector



SUCCESS:
EACH OF THE FOUR SENSORS OUTPUT IS IN LOGIC STATE "1"
(THEN TONE AND/OR GREEN LIGHT ARE AUTOMATICALLY TURNED
ON INDICATING TO OPERATOR THAT DEPTH POSITION AND PITCH
AND YAW ALIGNMENTS OF END EFFECTOR ARE OK FOR SUCCESSFUL
GRASP OF TARGET)

Figure 17. Analog drive logic for "success display"

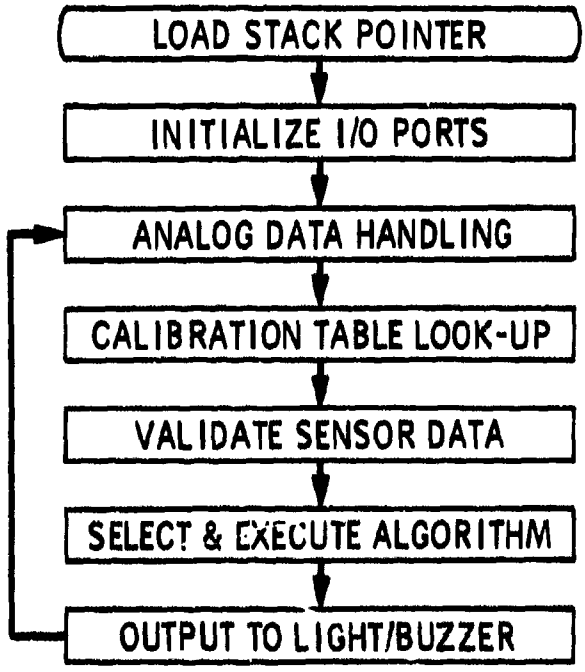
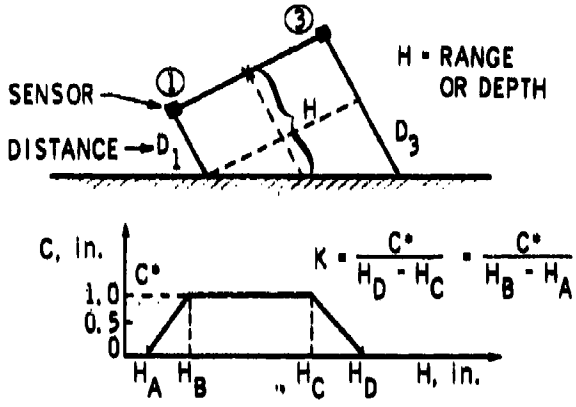


Figure 18. Computer program organization of digital drive for "success display"

MEASUREMENTS



C IS A MEASURE FOR PITCH AND YAW ERRORS; C = f(H)

H_A, H_B, H_C, H_D AND K (AND IMPLICITLY ALSO C*) ARE PRESET CONSTANTS

f(H) IS GIVEN BY THE TRAPEZOID FORMULA SHOWN ABOVE

SIMPLE ALGORITHM

① $A \leq D_1 \leq B$
② $A \leq D_2 \leq B$
③ $A \leq D_3 \leq B$
④ $A \leq D_4 \leq B$
⑤ $|D_1 - D_3| \leq C$
⑥ $|D_2 - D_4| \leq C$

IF ALL SIX CONDITIONS ARE TRUE THEN LIGHT/BUZZER ARE ON, OTHERWISE OFF

A, B, C ARE PRESET CONSTANTS

CONIC ALGORITHM

① $H_A \leq H \leq H_D$
WHERE $H = 1/2 (D_1 + D_3)$
$1/2 (D_2 + D_4)$

② $(D_1 - D_3)^2 + (D_2 - D_4)^2 \leq L$
WHERE $L = C^2 = [f(H)]^2$

IF BOTH CONDITIONS ARE TRUE THEN LIGHT/BUZZER ARE ON, OTHERWISE OFF

EXPANDED ALGORITHM

① $H_A \leq H \leq H_D$
WHERE $H = 1/2 (D_1 + D_3)$
$= 1/2 (D_2 + D_4)$

② $|D_1 - D_3| \leq C = f(H)$

③ $|D_2 - D_4| \leq C = f(H)$

IF ALL THREE CONDITIONS ARE TRUE THEN LIGHT/BUZZER ARE ON, OTHERWISE OFF

Figure 19. Success algorithms

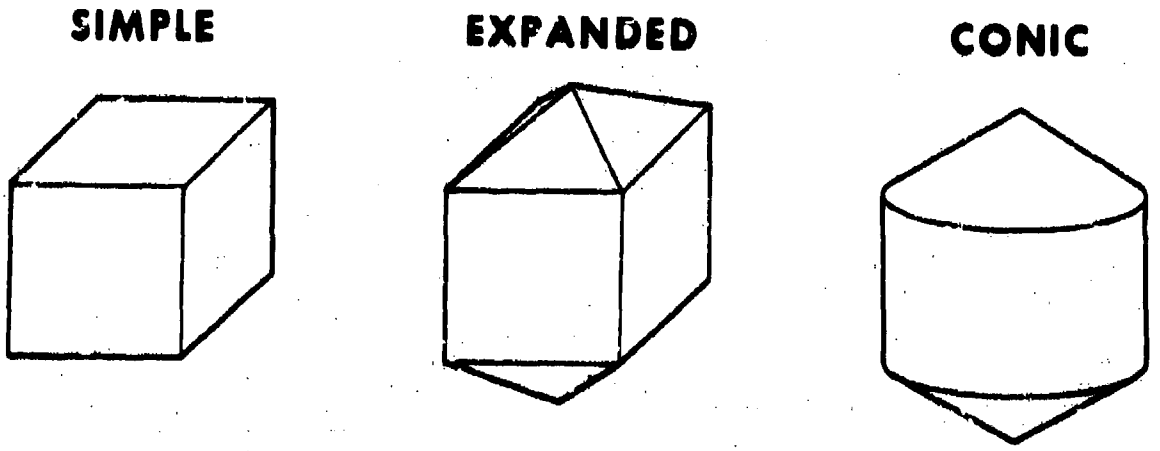


Figure 20. Successful grasp volumes corresponding to algorithmic definitions

The transfer of control and guidance information to the pilot through the manipulator forces.

by

Prof. Dr.-Ing. K.H. Doetsch and Dr.-Ing. W. Röger
Sonderforschungsbereich Flugführung
Techn. University Braunschweig
3300 Braunschweig, Postbox 3329
Federal Republic of Germany

ABSTRACT

Results are discussed of extensive investigations on a fixed cockpit simulator which aim at improving the control efficiency of the pilot. Information on the deviation from the desired or commanded flight path is transferred to the pilot by means of changes of control wheel forces proportional to, but in direction opposite to the control force required to reduce the deviation. Identical instrument display information is used simultaneously in some of the test series.

Quite remarkable improvements in flight path accuracy as well as reduced mental and physical workload are achieved, and the frequency range of disturbances that can effectively be corrected is extended. The essential cause of this phenomenon is identified as a rather drastic reduction in the human operator time delay when this particular type of kinesthetic cue is used. This enables the pilot to use higher gain in the control loop.

LIST OF SYMBOLS

$e(t)$	tracking error
e_{eff}	rms value of error
F	force transducer
$F_{MP}(s)$	model pilot describing function
FZS	Force Zero Shift
K_p	model pilot gain factor
t	time
T_I	human delay time constant
T_L	human lead time constant
T_N	neuro-muscular delay time constant
T_S	time constant of plant
V_S	gain of plant
$w(t)$	forcing function
w_{eff}	rms value of forcing function
$x(t)$	system output
$y(t)$	control signal (pilot's applied force)

y_{eff}	rms value of control signal
$\overline{\phi}_{ee}$	average power spectral density of tracking error, e
ϕ_{ww}	power spectral density of forcing function
ρ_{lin}	linear-coherence coefficient
σ	Force Zero Shift gain, degrees deflection of control column per unit tracking error, e
τ	human operator delay time
φ	phase margin
ω	frequency
ω_c	crossover frequency
ω_i	cut-off frequency in forcing function

INTRODUCTION

In the Sonderforschungsbereich Flugführung of the Braunschweig Technological University, we are engaged in a major research project on the optimal symbiosis of the human operator with automatic control of aircraft [1]. In the process of this research, the question arose whether airline pilots would, in the future, be content, within their task of flight path management, to select different fully automatic control modes or to perform flight path adjustments by pure press button operation, or whether they would prefer to "stay in the loop" by flying the aircraft through the AFCS in a control wheel steering fashion and thus be better adjusted to taking over instantaneously in the case of an irregularity.

If this latter approach were adopted, particularly during maneouvring flight in the terminal area and during landing, then improved means of communication between the aircraft dynamics and the pilot's sensory faculties would be most desirable. One way of achieving this was considered to exist in the use of kinesthetic cues via the manipulator,as suggested in different forms by Noggle [2] and Merhav [3].

In our particular experiments [4],we first used the control column in our existing flight simulator (Fig. 1) to provide kinesthetic cues to the operator. In the present flight development stage, a miniaturised manipulator, better adapted for the task is being used.

THE ACTIVE MANIPULATOR

The general idea of creating an appropriate kinesthetic cue may be explained by means of Fig. 2, which compares the principle of a conventional control system, including an artificial spring feel and force trim as used with most hydraulically operated aircraft control surfaces, with the principle of our "Active Manipulator". This manipulator, as far as the pilot is concerned, is conventional with typical steady state force-displacement characteristics. However, the zero force trim point is shifted continuously and automatically by any selected aircraft state parameter, and it is the force, F, exerted by the pilot on the control column rather than its deflection which is transmitted as the command

to the control system. In fact the mechanical control linkage is eliminated and therefore the system does not respond to manipulator position or deflection, nor does the manipulator move in sympathy with the control surfaces or the plant output. It does move, however, even if no force is applied by the pilot, in following the Force Zero Shift introduced by the motion of the spring feel system.

This Force Zero Shift (FZS) is defined in terms of the deflection of the control column or stick, measured in degrees (σ^0) per unit deviation, e, from the commanded flight path parameter. Fig. 3 demonstrates how this principle is translated into reality by means of an electro-hydraulic servo arrangement. The bottom of the control column is mechanically connected to a hydraulic jack. This receives two kinds of input signals, one from the load cell, measuring the pilot's applied force and providing the spring feel effect by equating this force input signal to a deflection feed back signal, and a second input signal, derived from a plant output parameter, causing the Force Zero Shift of the spring feel and thus moving the free column position.

This shift is arranged in such a direction that , if the pilot attempts to hold the column in its original position, he has to apply an opposing force, i.e. a force in such a sense that the load cell signal tends to reduce the plant output deviation, just as an unaided control force input in the same direction would do. This direction of the Force Zero Shift is, nota bene, opposite to an ordinary trim servo input aimed at reducing the same deviation.

This is to provoke, by the kinesthetic cue of the Force Zero Shift, a reflex type reaction of the pilot in the correct phase with the required control input. It relieves him of a conscious decision making process to find this correct phasing, which would consume much time and still be apt to produce erroneous inputs.

The magnitude of the kinesthetic cue can easily be adjusted by changing the gain setting of the Force Zero Shift (see the block "Gain Variation" in Fig.4). The gain of the Force Zero Shift response to aircraft state variations is obviously one of the main parameters of our investigation. It can be conveniently expressed in terms of zero force angular control column deflection, σ, as mentioned before, once a suitable feel spring stiffness has been established and fixed.

The pilot is free to react to the force cue with any amplification he may choose, i.e. with his personal linear gain factors, K_p. If, in particular, he follows the strategy of doing no more nor less than holding the control column fixed in its initial position, then the unique gain factor needed to achieve this may be referred to as "Reference Gain". It is worth noting that it is only dependent on constant settings of the system.

EXPERIMENT

As evident from Fig. 4, this basic experiment was conducted as a single loop manual compensatory tracking task in the aircraft pitch degree of freedom. The aircraft behaviour was reduced to the Laplace form

$$x(s) = \frac{V_s}{s(T_s s+1)} \cdot y(s).$$

In most of the test runs, the operator received, in addition to the kinesthetic cue via the Active Manipulator, identical visual information of the tracking error through a modified artificial horizon display. The forcing function, w(t), was composed of ten harmonic sine-waves at different, discrete frequencies, taken from AFFLR-TR-65-15 by McRuer et al., (Fig. 14). For the assessment of the ability of the operator to follow the forcing function, both, the rms value of the error, e_{eff}, and, as a measure of his control activity, the rms value of the control signal, y_{eff}, were evaluated. For each series of runs, the average operator transfer function was established over the whole frequency range. The linearity of the pilot describing function was judged by the ratio of the power of his control activity, y(t), in the forcing function frequencies to the total power. This ratio was defined as linear-coherence coefficient, ρ_{lin}. In addition, the power spectral densities of the tracking error were established for the forcing function frequencies in order to observe the effect of the Active Manipulator on the system performance, particularly with respect to an extension of the controllable frequency range.

Test runs of 4 minutes each were performed by 6 subjects and repeated 5 times for each setting of σ. Thus, 30 individual test values were combined to form one averaging point in the plotting of the test results.

TEST RESULTS

Fig. 5 shows what remarkable improvement in tracking performance is achieved even with moderate force cues. At a Force Zero Shift of $\sigma = 6^0$, the average tracking error is halved when compared with conventional control ($\sigma = 0$). The same order of improvement is evident in the intrasubject as well as in the intersubject standard deviations. This indicates that the operators are increasingly compelled to use a uniform and very efficient control strategy.

One might assume that these improvements are paid for by increased pilot's workload, in particular higher control activity. Fig. 6 shows that the opposite is true. The pilot's control effort is, for all Force Zero Shift gains, slightly reduced. And, incidentally, the intra- and intersubject variation are gradually eliminated with increasing gain.

The power spectral densities of the tracking error (Fig. 7) show that the improvements in tracking accuracy cover the whole frequency range that can be controlled manually by the human operator. In fact, the controllable range is extended, and enhanced control is achieved at the lower as well as at the upper range limit.

One reason for the improvement in tracking performance lies in the improved linearity of the operator's control activity (Fig. 8). Linearity tends, with increasing gain of the Force Zero Shift, towards unity, i.e. the pilot does not waste much energy on irrelevant frequencies.

This high degree of linearity permits a consideration of the pilot's measured activity in terms of the parameters of the well known pilot describing function in its purely linear form (Fig. 9). The changes in some of the parameters are quite remarkable. For instance, the pilot's gain, K_p, related to that with conventional control, more than doubles its value at the moderate FZS gain of $\sigma = 6^0$. It does not, at this FZS gain, limit itself to the Reference Gain which would result from just blocking the control column motion by the appropriate

reaction force. The pilot obviously adjusts his K_p to suit the task. In fact, it is to be noted that the test subjects always tried to increase K_p to such a value that the remaining phase margin judged to be necessary for system stability stayed at about 40^o.

The most remarkable change occurs in the time delay parameter, τ, which drops drastically to an imperceptibly low value at $\sigma = 6^o$ or more (note the sensitivity of the scaling in the ordinate of Fig. 9.). The neuromuscular time constant, T_N, first rises slightly but decreases once τ has gone to zero, and it disappears at a gain setting of $\sigma = 12^o$ or more. The effect of this near-elimination of the human reaction delay times manifests itself in the much improved phasing of the pilot's control reactions to the demands of the forcing function.
In Fig. 10, the pilot's open loop crossover frequency and the phase margin are plotted against the FZS gain parameter. This emphazises what has just been said: The human operator utilizes the phase advance gained through the reduction of τ and T_N in order to increase his gain factor, K_p, until the remaining phase margin puts a limit on further increases. The corresponding crossover frequency, ω_c, monotonically increases with the FZS gain. Thus, the Active Manipulator permits the operator to correct tracking errors more rapidly and, consequently, to increase tracking accuracy.

Another interesting aspect of the Active Manipulator is shown in Fig. 11. During the whole period of an unskilled operator's training phase, the tracking accuracy is twice as good with a Force Zero Shift gain of $\sigma = 6^o$ as with a conventional control ($\sigma = 0$), and the initial learning phase compares even better; the operator adapts more quickly to the control task. In fact, a constant level of performance is reached after only 10 test runs compared with 20 for conventional control. In Fig. 12, it becomes obvious that the tracking performance is much nearer the linear optimum right from the start.

The essential point of any explanation for this and for the positive results shown earlier is the fact that, with the kinesthetic cues of the Active Manipulator, there is neither a chance for an error nor a reason for hesitation to apply the control input in the correct direction. To put it in a simplified form:

The pilot is compelled, by instinctive reaction to the force cue, to give a control input in the required phase. After that, it only remains for him to adjust the amplitude of his response to the magnitude of the stimulus.

FURTHER STUDIES

In order to obtain a first indication on the applicability of the Active Manipulator in real life control tasks, two further steps in the simulator investigation were taken. The plant characteristics were modified with respect to gain and time constant. Increases in gain showed a much improved linearity throughout the pilot's responses, whereas that of the conventional control deteriorated considerably with increasing plant gain. The time constant of the 1st order plant was varied from zero to 1.5 seconds. The linearity deteriorated only slightly from 0.98 to 0.92, but the average rms tracking error (Fig. 13) increased considerably, as would be expected. This can be cured by shaping of $\sigma(e)$. But even without this, the deterioration was much less than that for conventional control.

First flight tests were performed on a Dornier Skyservant aircraft with command control (Figs. 15 and 16). Stick force fore and aft commanded vertical acceleration, $\dot{h}$, and sideways force commanded yaw acceleration, $\ddot{\psi}$. Stepwise changes of heading manoeuvres were performed with and without the Active Manipulator. Control was remarkably improved with respect to heading accuracy as well as control smoothness when the Active Manipulator was employed. The same striking improvement was observed with stepwise changes in rate of descent.

We are confident that the Active Manipulator will prove an excellent means to enable the pilot to "stay in the loop" with largely automated flight control. He can more effectively control the aircraft, providing that a source of aircraft state information is available for the Force Zero Shift.

REFERENCES

|1| Doetsch, K.H., The proper symbiosis of the human pilot and automatic flight control, Eighteenth Lanchester Memorial Lecture, The Aeronautical Journal, June 1975

|2| Noggle, P.L., Manual control of unstable vehicles using kinesthetic cues, Massachusetts Institute of Technology, MV-69-4

|3| Merhav, S.J., On optimality in human control tasks,10th Congress of the International Council of the Aeronautical Sciences, 1976

|4| Röger, W., Das Bedienelement als Informationsträger bei Bahnführungsaufgaben, Dr.-Ing. Thesis, Institut für Flugführung, TU Braunschweig, December 1978

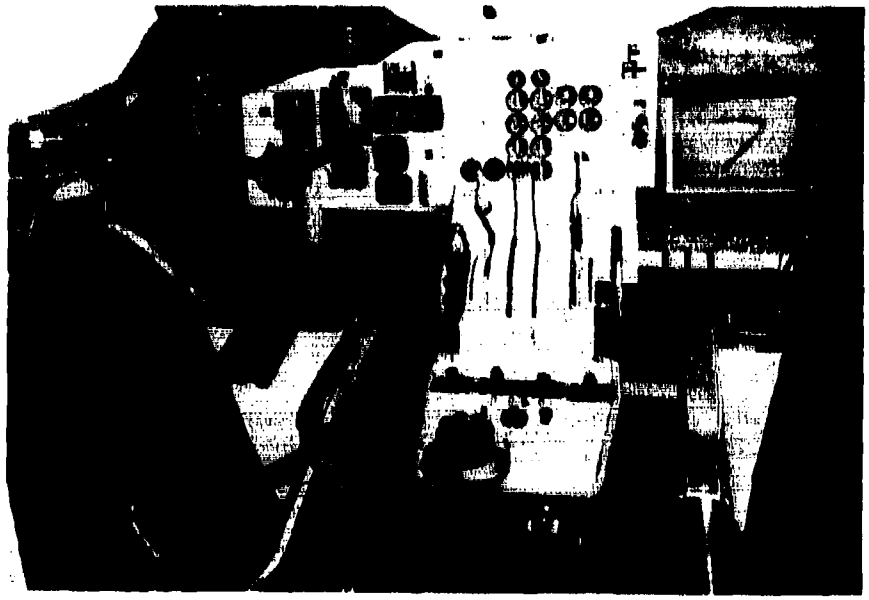


Fig. 1 Simulator Cockpit

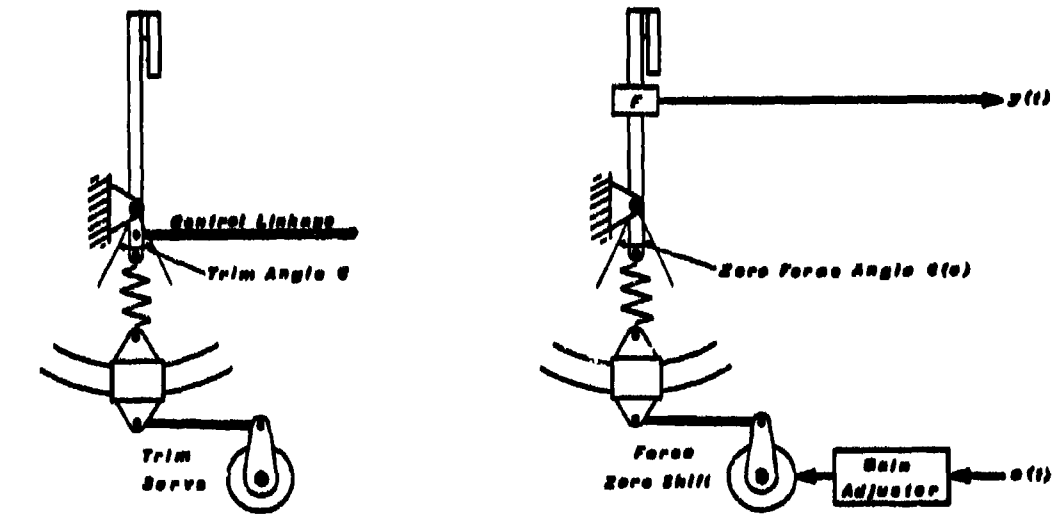


Fig. 2 Control Column with Mechanical Feel System

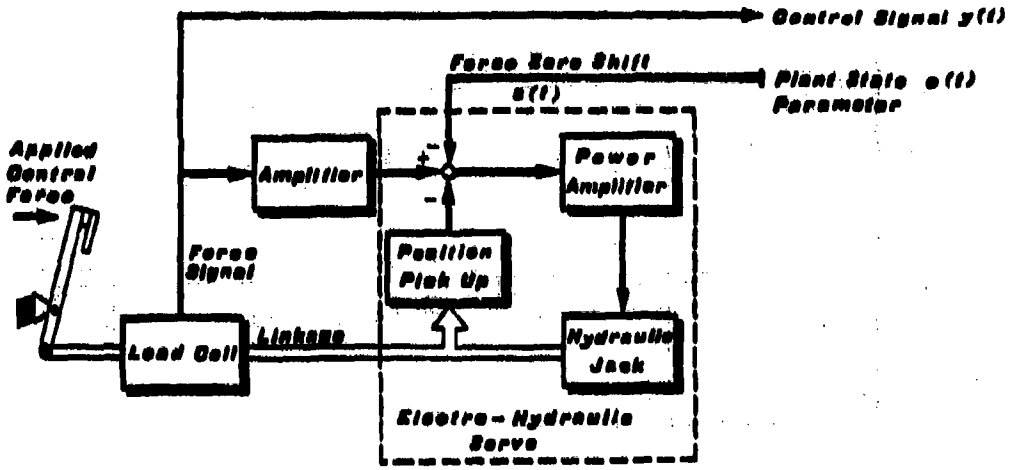


Fig. 3 Artifical Feel System with Force Zero Shift

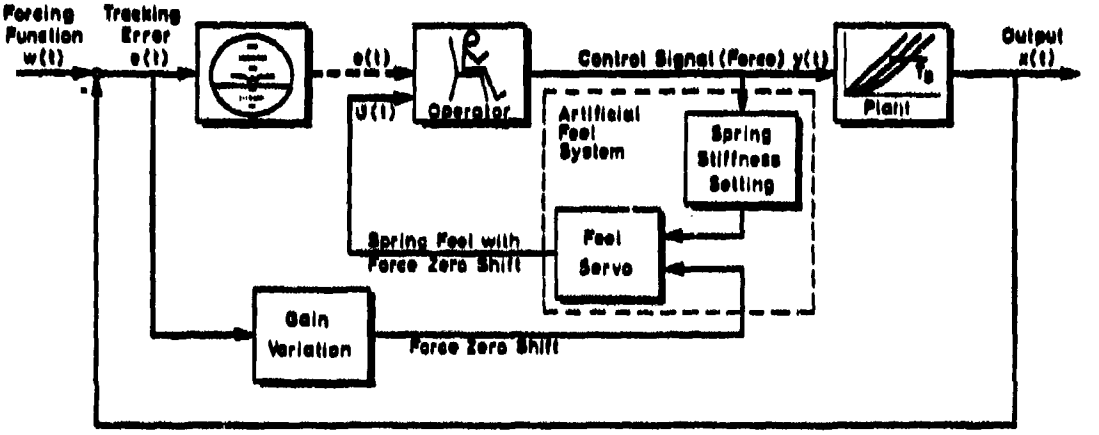


Fig. 4 Simulator Test Arrangement

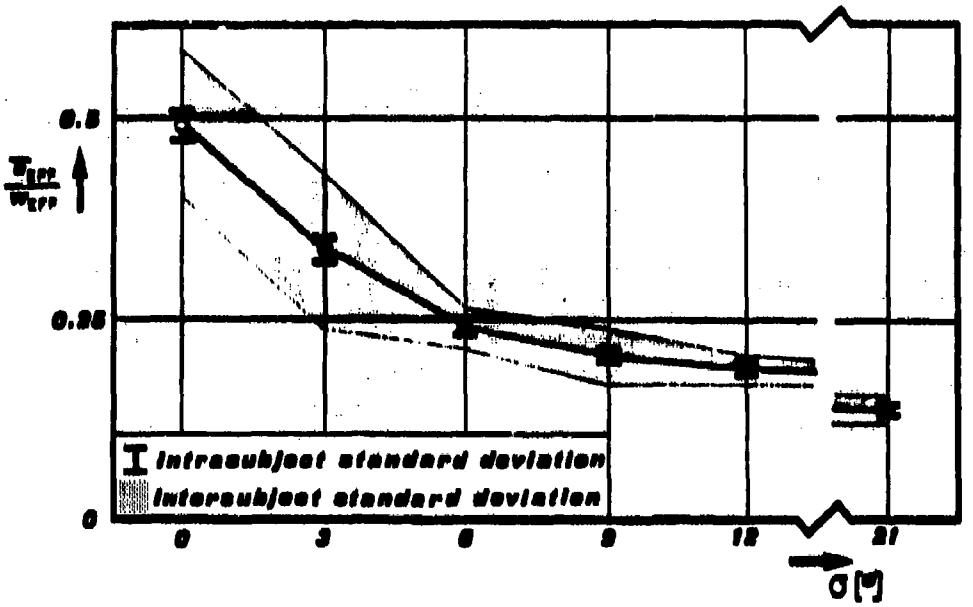


Fig. 5 Average Relative Tracking Error

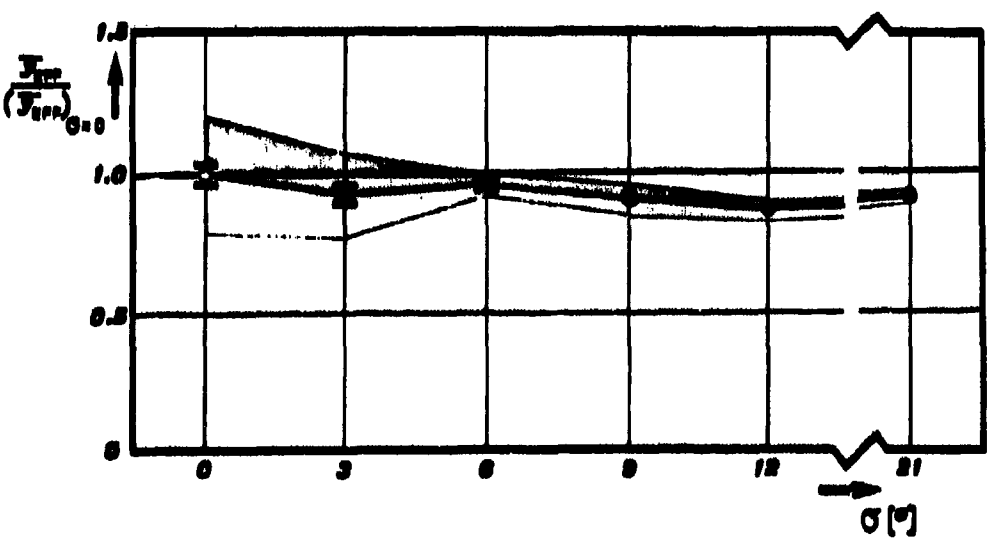


Fig. 6 Average Control Effort

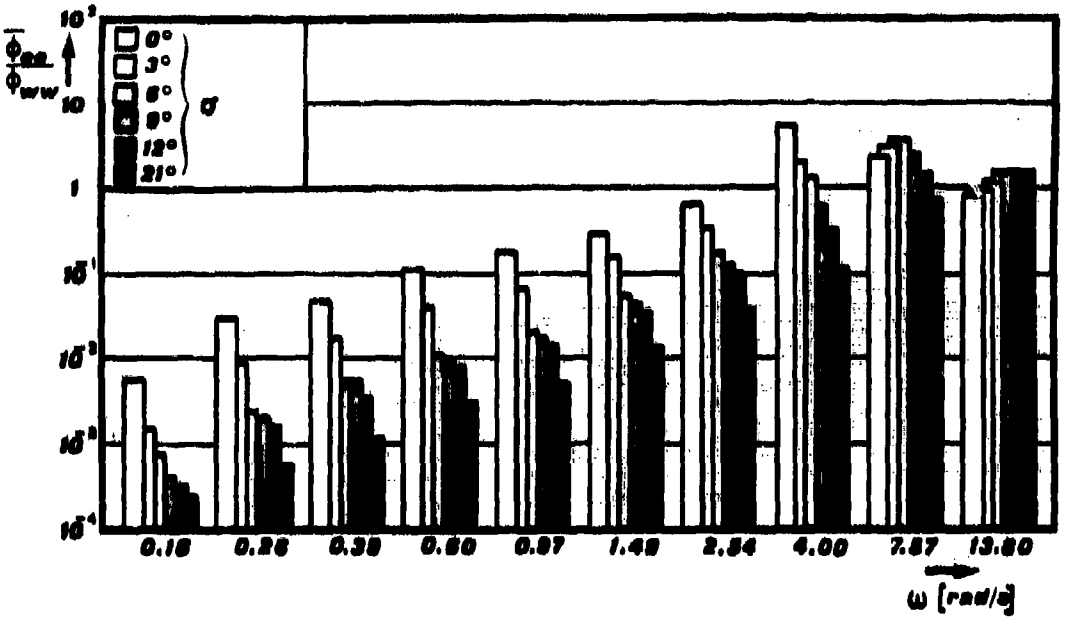


Fig. 7 Average Power Spectral Density of Tracking Error

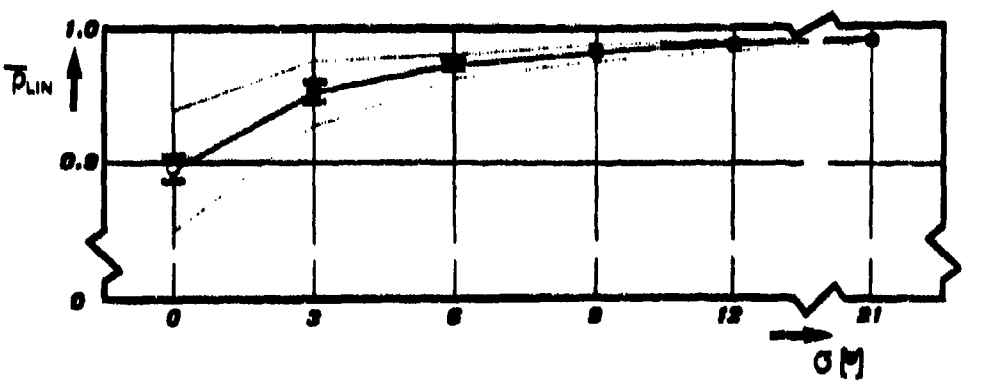


Fig. 8 Average Linear-Coherence Coefficient

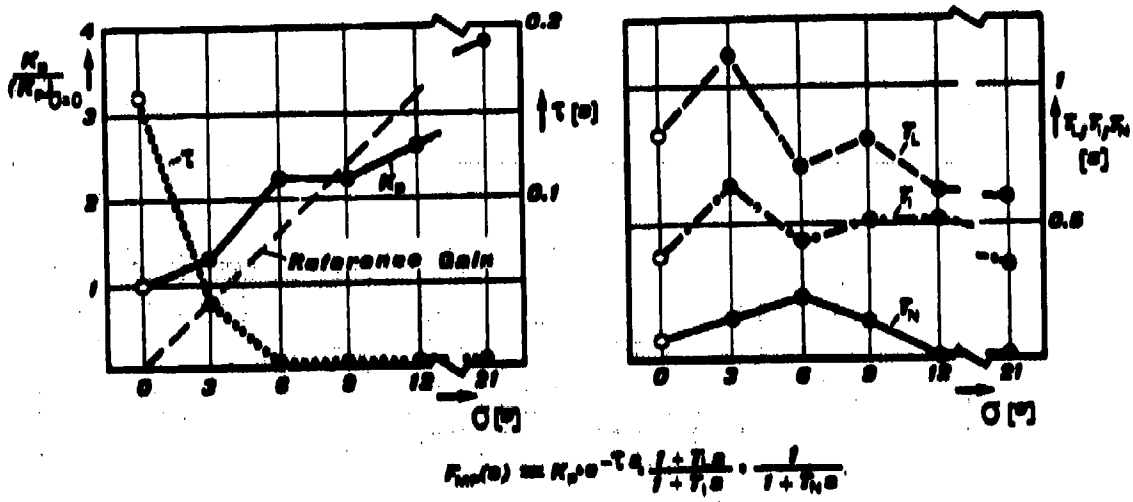


Fig. 9 Parameters of the Model Pilot Describing Function

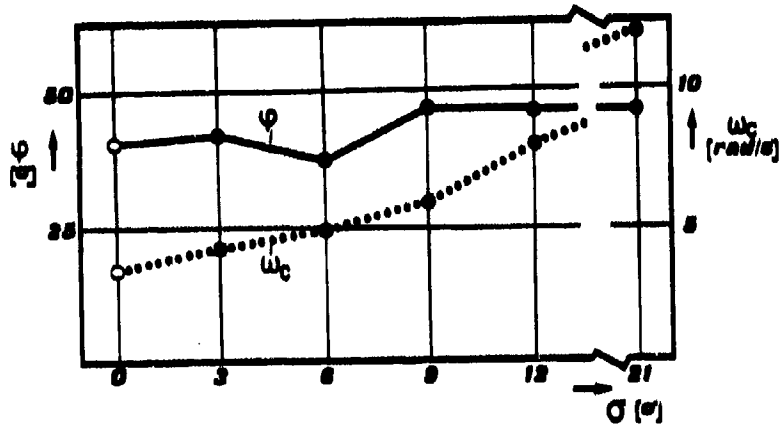


Fig. 10 Phase Margin and Crossover Frequency

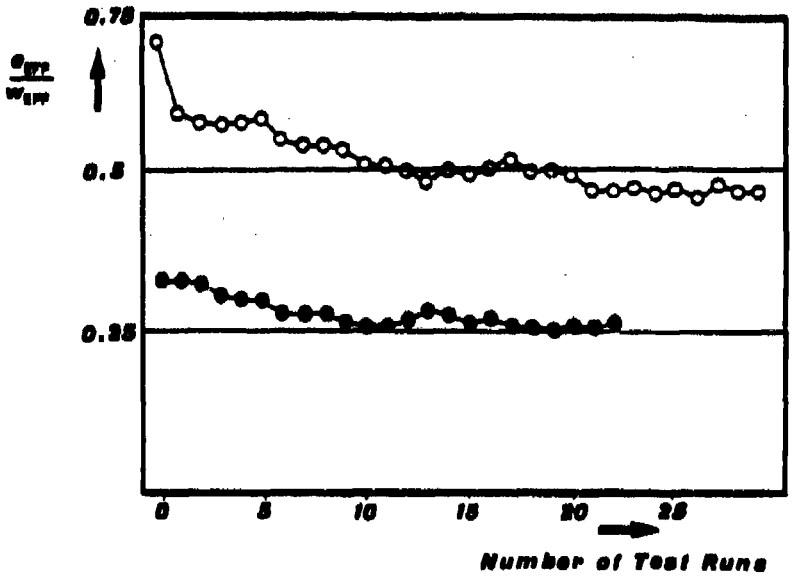


Fig. 11 Rel. Tracking Error during Training Phase

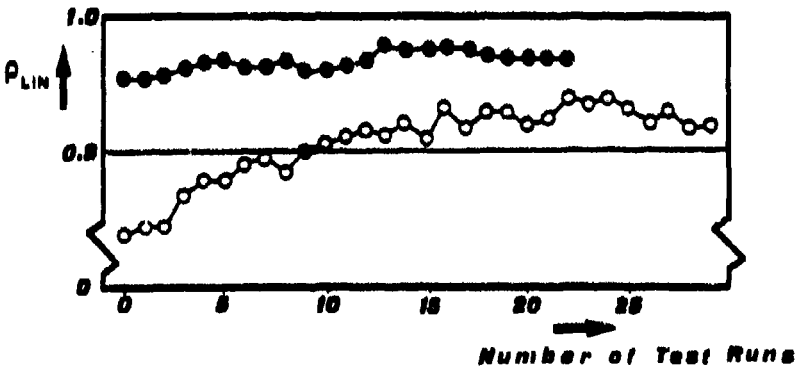


Fig. 12 Linear-Coherence Coefficient during Training Phase

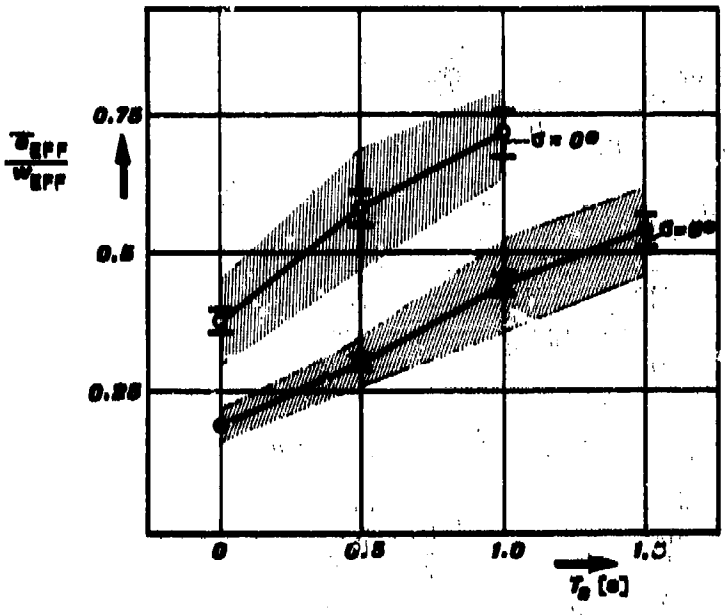


Fig. 13 Effect of Time Constant of a 1st Order Plant on Tracking Error

k	1	2	3	4	5	6	7	8	9	10
ω_k [rad/s]	0.157	0.262	0.393	0.602	0.989	1.49	2.54	4.0	7.57	13.80
$\lambda = \frac{\omega_k}{\omega_0}$	6	10	15	23	37	57	97	154	289	527
A_k	1	1	1	1	1	1/0.1	1/0.1	0.1	0.1	0.1

$$w(t) = \sum_{k=1}^{10} A_k \sin \omega_k \cdot t$$

Fig. 14 Frequencies, Number of Periods and Amplitudes for the Components of the Forcing Function

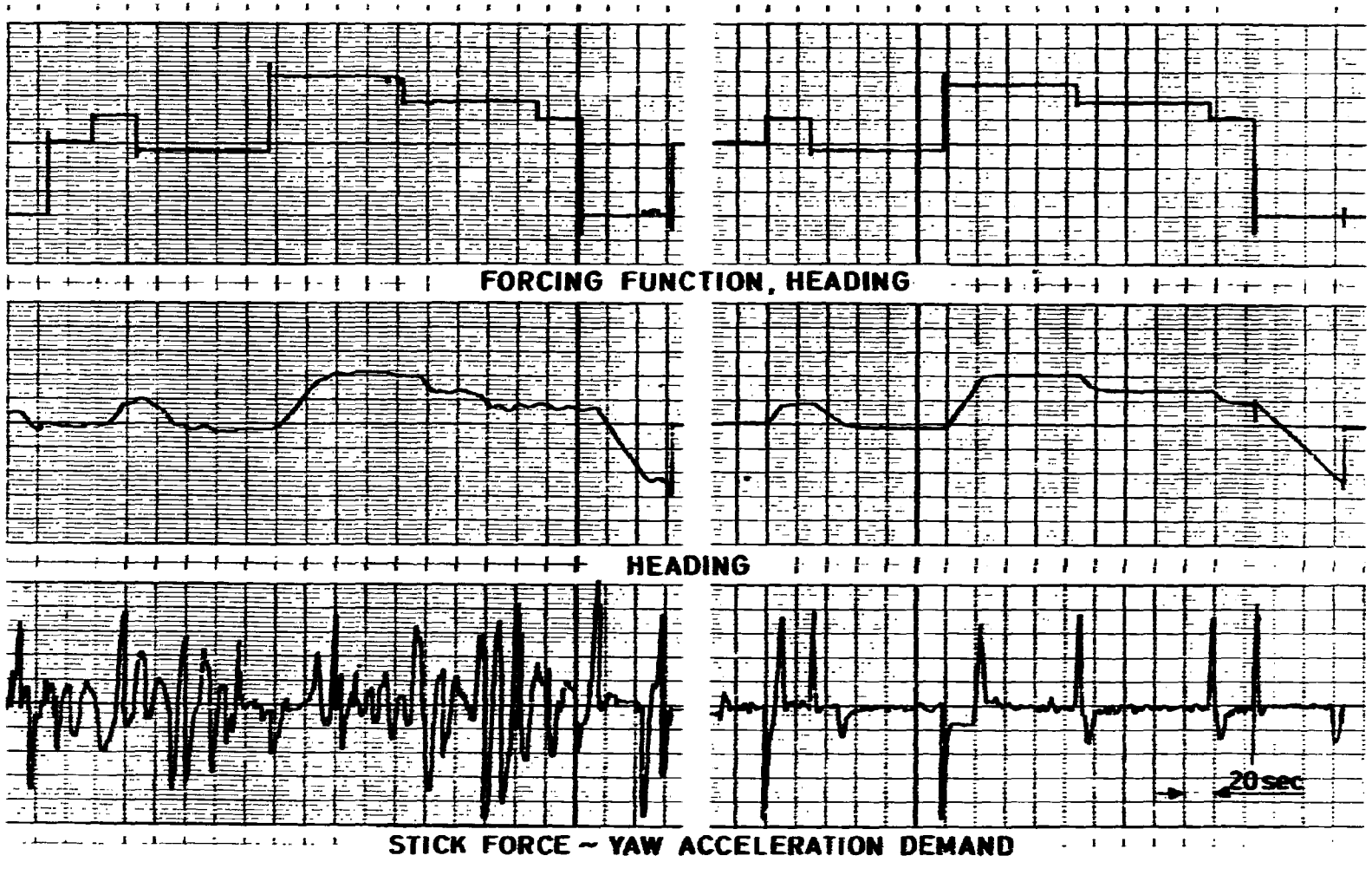


Fig. 15 Flight Test

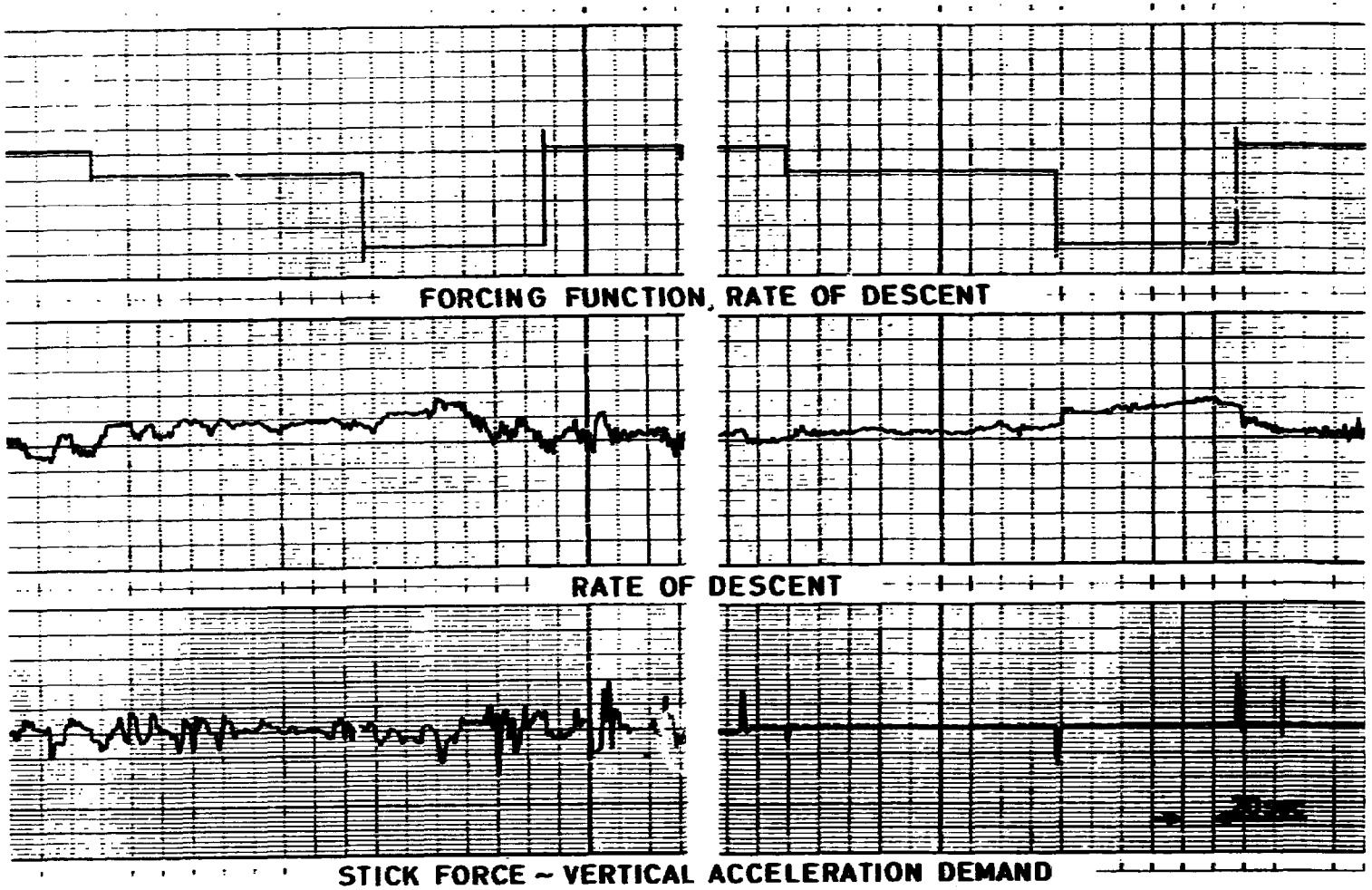


Fig. 16 Flight Test

CONTROL/DISPLAY SYSTEM SYNTHESIS USING CLOSED LOOP PERFORMANCE CRITERIA AND PILOT MODELING

John E. Hart
Flight Sciences Division
Lockheed-Georgia Company
Marietta, Georgia 30063

ABSTRACT

A sufficient condition for good aircraft handling qualities is that no pilot equalization is required to obtain the desired response. The "desired response" is defined in terms of the closed-loop control characteristics.

For a given aircraft, and assuming a priori that the pilot does not introduce equalization, the control/display system requirements may be obtained by extracting the aircraft and pilot models from the model of the desired closed-loop performance. Control and display system parametric control laws have been examined by this technique for lightly damped and unstable aircraft. The use of compensation, as well as conventional augmentation, has been considered with interesting ramifications.

INTRODUCTION

Most of the handling quality criteria of the past have been based on experience which indicates that good handling qualities result when the characteristics of one of the elements of the control loop, the aircraft itself, are within certain bounds. If these characteristics are not inherent in the aircraft design, the approach has usually been to augment them with a conventional stability augmentation system. However, all manual control systems are in reality closed-loop controls with the control loop consisting of the (attitude error) display, the pilot, the forward loop control elements and the aircraft (with or without stability augmentation). Therefore, it is plausible to introduce compensating characteristics into the display or forward loop control to provide desirable handling qualities, since "the pilot evaluates the total response of the airplane to his inputs and is not concerned with, or aware of, the characteristics of the individual elements which combine to produce that response" (Reference 1).

Results of past handling quality criteria investigations and pilot modeling studies have shown that a sufficient condition for satisfactory handling qualities is that the pilot is not required to introduce dynamics into the control loop (References 1 and 2). For a given aircraft, and assuming a priori that the pilot does not introduce equalization, requirements for various display or control compensators, augmentation schemes, or combinations, may be derived from known desirable closed-loop characteristics. This technique was used by

Chen and Boothe in designing augmentation systems for fighter aircraft (Reference 3), but previous applications to the design of compensators are not known.

CLOSED-LOOP CRITERIA

Closed-loop performance, which results in satisfactory handling qualities, has been defined in terms of the closed-loop frequency response (References 2 and 4). Basically, the requirement is that the frequency response should be flat from zero frequency to some bandwidth, BW, value. The value of BW depends upon the aircraft; it tends to be about 3 radians per second for light, highly maneuverable aircraft and about 2 radians per second for heavy aircraft. Low-frequency droop, where the amplitude ratio drops below three decibels, is not acceptable. A bandwidth of less than BW is not acceptable. If a resonant rise in the frequency response occurs, it is undesirable that it exceeds 3 decibels and unacceptable for it to exceed about 10 decibels. These requirements are generally consistent with the closed-loop "criteria" given by McRuer in Reference 1.

PILOT ADAPTATION

The classical pilot model is made up of three basically independent parts. The first is his reaction time delay (and neuromuscular lags) over which he has no adjustment capability; the second is the equalization which he adapts depending upon the dynamics of the remainder of the control loop; and the third is his gain. In what follows it is assumed that the dynamics of the control loop are such that the pilot has no need to introduce equalizing dynamics, and therefore, he does not. It is assumed, however, that he adjusts his gain to achieve the desired bandwidth.

THE ANALYSIS METHOD

Consider a statically unstable aircraft with the pitch angle, θ, controlled by the elevator, δ_e. The transfer function may be represented by:

$$-\frac{\theta}{\delta_e} = \frac{K_A}{S(T_A S - 1)}$$

Neglecting the pilot's reaction time delay for the moment, the pilot is represented as a simple gain, K_p. This gain is the amount of elevator introduced per unit error between the desired and actual pitch angles. The block diagram for this attitude control loop is shown in Figure 1.

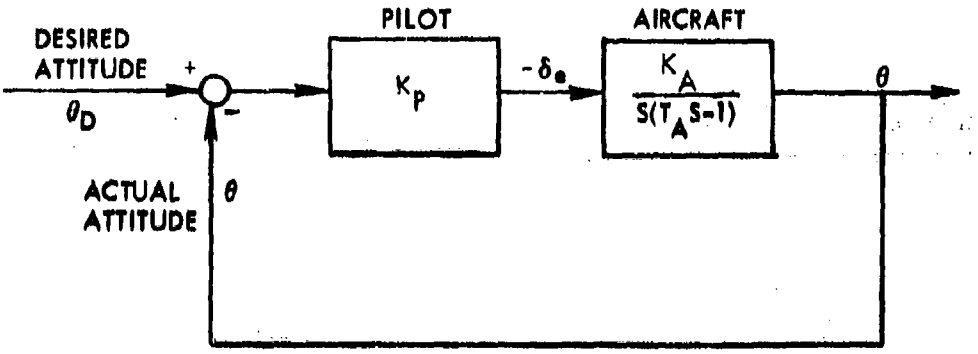


Figure 1. Piloted Attitude Control Loop - Unstable Aircraft

The closed-loop transfer function, relating the actual attitude to the desired attitude is:

$$\frac{\theta}{\theta_D} = \frac{1}{\frac{T_A}{K_p K_A} S^2 - \frac{1}{K_p K_A} S + 1}$$

This is re-written in the form:

$$\frac{\theta}{\theta_D} = \frac{1}{\frac{S^2}{\omega^2} + \frac{2\zeta}{\omega} S + 1}$$

where the natural frequency, ω, is:

$$\omega = \sqrt{\frac{K_p K_A}{T_A}}$$

and the damping, ζ, is:

$$\zeta = -\frac{1}{2\omega T_A}$$

By adjusting his gain the pilot can maintain the desired closed-loop natural frequency, but the damping is negative at all frequencies. The closed-loop criteria can not be met. Therefore, the assumption that the pilot acts as a simple gain is invalid in this case and the sufficient condition for acceptable handling qualities is not met.

The Addition of SAS

Consider the same aircraft but with the addition of a pitch rate feedback SAS. The closed-loop attitude control is shown in block diagram form in Figure 2. The transfer function relating actual attitude to desired attitude is:

$$\frac{\theta}{\theta_D} = \frac{1}{\frac{T_A}{K_p K_A} s^2 + \left(\frac{K_F}{K_p} - \frac{1}{K_p K_A}\right) s + 1}$$

or

$$\frac{\theta}{\theta_D} = \frac{1}{\frac{s^2}{\omega^2} + \frac{2\zeta}{\omega} s + 1}$$

where

$$\omega = \sqrt{\frac{K_p K_A}{T_A}}$$

as before, but the damping is now:

$$\zeta = \frac{K_F K_A}{2\omega T_A} - \frac{1}{2\omega T_A}.$$

If K_F, the gain of the pitch rate feedback, is made sufficiently large the damping may be set at any desired positive value and acceptable handling qualities attained by satisfying the closed-loop criteria. Note, however, that the positive part of the damping expression includes the aircraft gain parameter, K_A. This gain varies with airspeed. To keep the range of damping within acceptable limits over the flight envelope, it is necessary to vary the SAS gain, K_F, as a function of airspeed. This gain-varying feature obviously adds to the system complexity.

Control Compensation

Consider once again the same unstable aircraft but with the addition of a compensator between the pilot and the elevator deflection (and without SAS.) A compensator which produces an elevator deflection equal to the pilot's input plus a deflection proportional to the rate of change of the pilot's input is shown in the block diagram for this system, Figure 3.

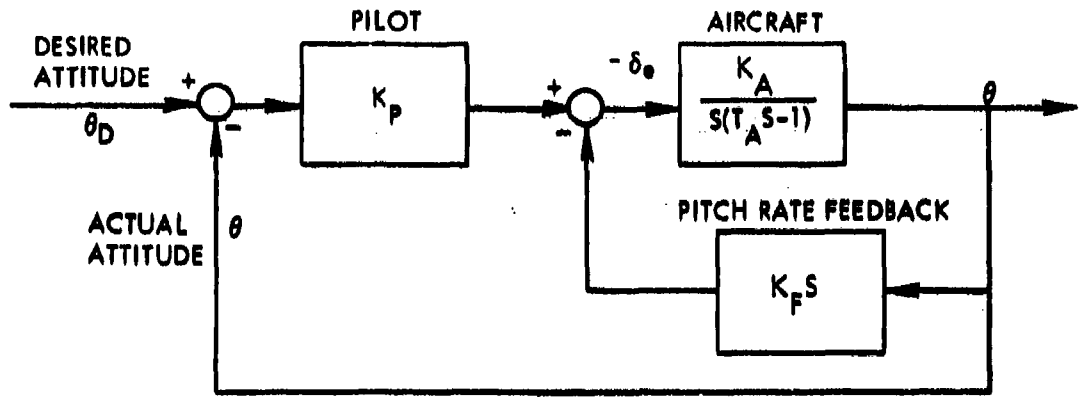


Figure 2. Unstable Aircraft With SAS

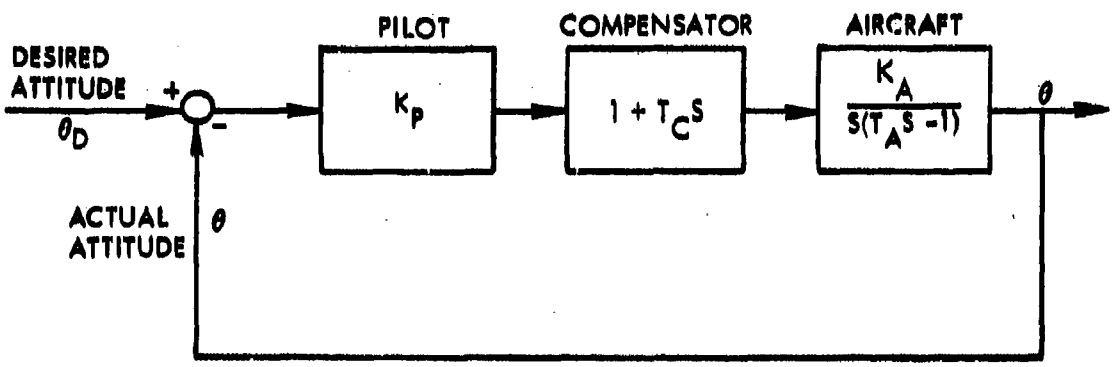


Figure 3. Unstable Aircraft With Control Compensation

The closed-loop transfer function is:

$$\frac{\theta}{\theta_D} = \frac{T_C S + 1}{\frac{T_A}{K_P K_A} S^2 + \left(T_C - \frac{1}{K_P K_A}\right) S + 1}$$

Again changing the form of the denominator

$$\frac{\theta}{\theta_D} = \frac{T_C S + 1}{\frac{S^2}{\omega^2} + \frac{2\zeta}{\omega} S + 1}$$

where, once more:

$$\omega = \sqrt{\frac{K_p K_A}{T_A}}$$

but

$$\zeta = \frac{\omega T_C}{2} - \frac{1}{2\omega T_A}$$

Acceptable closed-loop damping values can usually be attained by selecting a large enough value of the compensator parameter, T_C. It is important to note that the positive term in this damping expression does not contain the aircraft gain parameter, K_A, as is the case if damping is provided by a SAS.

A condition which limits the use of a control compensator of the form shown is that the resulting lead term in the numerator of the closed-loop transfer function will cause a rise in the frequency response which may be unacceptable if T_C is too large. Thus, T_C must be large enough to provide the desired damping but not so large as to produce an unacceptable rise in the frequency response. The efficacy of this type of compensation, therefore, is reduced as T_A decreases.

Display Compensation

With the same unstable aircraft as before, the addition of a pitch rate command display compensator can produce acceptable closed-loop characteristics. A means is provided to allow the pilot to match the pitch rate to the difference between the desired attitude and the actual attitude. His control action, in this case, is proportional to the error in this matching process. Figure 4 is a block diagram of the control loop, where it is again assumed that good handling qualities are achieved and the pilot, therefore, acts as a simple gain.

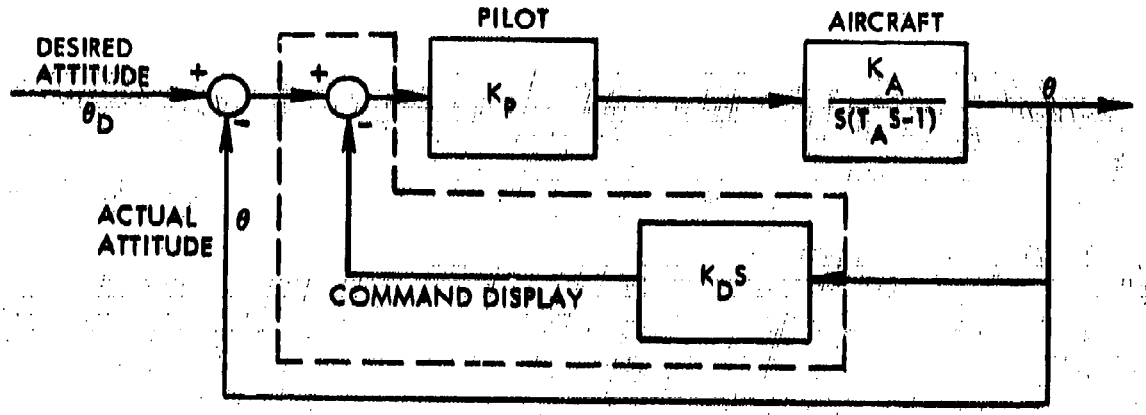


Figure 4. Unstable Aircraft With Display Compensation

The closed-loop transfer function is:

$$\frac{\theta}{\theta_D} = \frac{1}{\frac{T_A}{K_p K_A} S^2 + (K_D - \frac{1}{K_p K_A}) S + 1}$$

where, as before:

$$\omega = \sqrt{\frac{K_p K_A}{T_A}}$$

Damping in this case is:

$$\zeta = \frac{\omega K_D}{2} - \frac{1}{2 \omega T_A}$$

This expression is remarkable. Not only is the damping independent of the aircraft gain parameter, but because the numerator of the closed-loop transfer function is simply unity, there is no limit to the maximum value of the stabilizing parameter, K_D, in achieving the desired closed-loop damping.

CONSIDERATION OF PILOT REACTION TIME

The pilot's reaction time delay, omitted above to permit a less cluttered explanation of the analysis method, may be included by assuming the pilot's describing function to be:

$$K_p\, e^{-\tau S} = K_p\, \frac{-\frac{\tau}{2}S + 1}{\frac{\tau}{2}S + 1}$$

This more representative expression still assumes that the handling quality criteria for the closed-loop system are satisfactory and the pilot introduces no equalization. It is also still assumed that the pilot adjusts his gain to achieve the target bandwidth.

Control Compensation

Consider the control of Figure 3 but with the pilot represented by the describing function above. The closed-loop transfer function becomes:

$$\frac{\theta}{\theta_D} = \frac{(T_C S + 1)\,(-\frac{\tau}{2} S + 1)}{\frac{\tau T_C}{2 K_p K_A} S^3 + (\frac{T_A}{K_p K_A} - \frac{\tau}{2 K_p K_A} - \frac{\tau T_C}{2}) S^2 + (T_C - \frac{\tau}{2} - \frac{1}{K_p K_A}) S + 1}$$

Assume the form

$$\frac{\theta}{\theta_D} = \frac{(T_C S + 1)\,(-\frac{\tau}{2} S + 1)}{(TS + 1)\,(\frac{S^2}{\omega^2} + \frac{2\zeta}{\omega} S + 1)}$$

Parametric expressions for ω, ζ and T may be derived and the product $K_p K_A$ eliminated from them. Substituting known (References 2 and 3) numerical values for ω and τ and solving for damping as a function of the compensator time constant, T_C, for fixed values of the aircraft divergent time constant, T_A, results in a relatively simple quadratic form. For example, with

$\omega = 2$ radians/sec and $\tau = 0.3$ second

and with T_A assumed equal to 2 seconds:

$$\zeta = \frac{1.392 + T_C \mp \sqrt{4.698 + T_C^2 - 2.256\, T_C}}{2}$$

Plots of the closed-loop damping versus the compensator time constant are shown in Figure 5 for three values of T_A. Selecting a value for damping, the value of T may be obtained from the expression:

$$T = \frac{\frac{\tau}{2}\, T_A \left(\frac{\tau}{2} T_C\, \omega^2 + 1\right)}{T_A (1 - \tau\zeta\omega) - \frac{\tau}{2}}$$

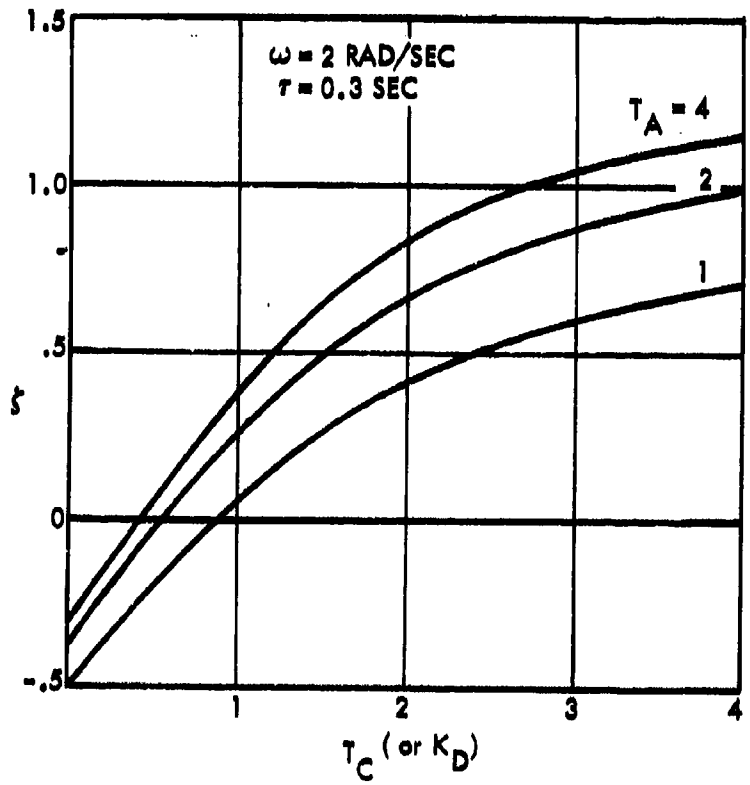


Figure 5. Damping of Quadratic Versus Compensating Parameter

Thus, a closed-loop frequency response may be plotted, as shown in the upper curve of Figure 6, for the parameter set:

T_A	= 2	assumed for aircraft
τ	= 0.3	assumed
ω	= 2	assumed (maintained by pilot gain adjustment)
T_C	= 1.15	selected
ζ	= 0.35	computed (see Figure 5)
T	= 0.79	computed

This frequency response, while considerably improved over the uncompensated case, does not result in satisfactory handling qualities, since the resonance exceeds three decibels. If the compensator time constant is increased, the numerator term tends to increase the resonance. On the other hand, a decrease in the compensator time constant tends to decrease the damping of the denominator quadratic which also increases the resonance.

Satisfactory handling qualities can be obtained with this type compensator if the aircraft divergent time constant is somewhat larger.

Display Compensation

Considering the system of Figure 4, but with the more realistic pilot model, a similar analysis may be performed. The closed-loop transfer function for this case is:

$$\frac{\theta}{\theta_D} = \frac{-\frac{\tau}{2} S + 1}{(T S + 1)\left(\frac{S^2}{\omega^2} + \frac{2\zeta}{\omega} S + 1\right)}$$

The expressions for damping and denominator time constant turn out to be identical to those for the control compensator case above but with the display gain K_D substituted for T_C. (Note that Figure 5 shows the ordinate as T_C or K_D.)

The lower frequency response in Figure 6 is for a display compensated case, where all other parameters are the same as the control compensated case. Satisfactory handling qualities are indicated because the closed-loop characteristics meet the criteria. Higher values of K_D would afford satisfactory handling qualities with even lower values of the aircraft divergent time constant.

CAVEAT

The parametric analysis method is useful for synthesizing control systems in general and compensation techniques in particular. Insight into the effects of various parameters and the sensitivity of the response to parametric variations or tolerances is provided. However, care must be taken to consider realistic limitations such as limiting the pilot gain to acceptable values, and control surface deflection and rate limits. As the aircraft and system models become more exact, and involved, parametric analyses become extremely tedious.

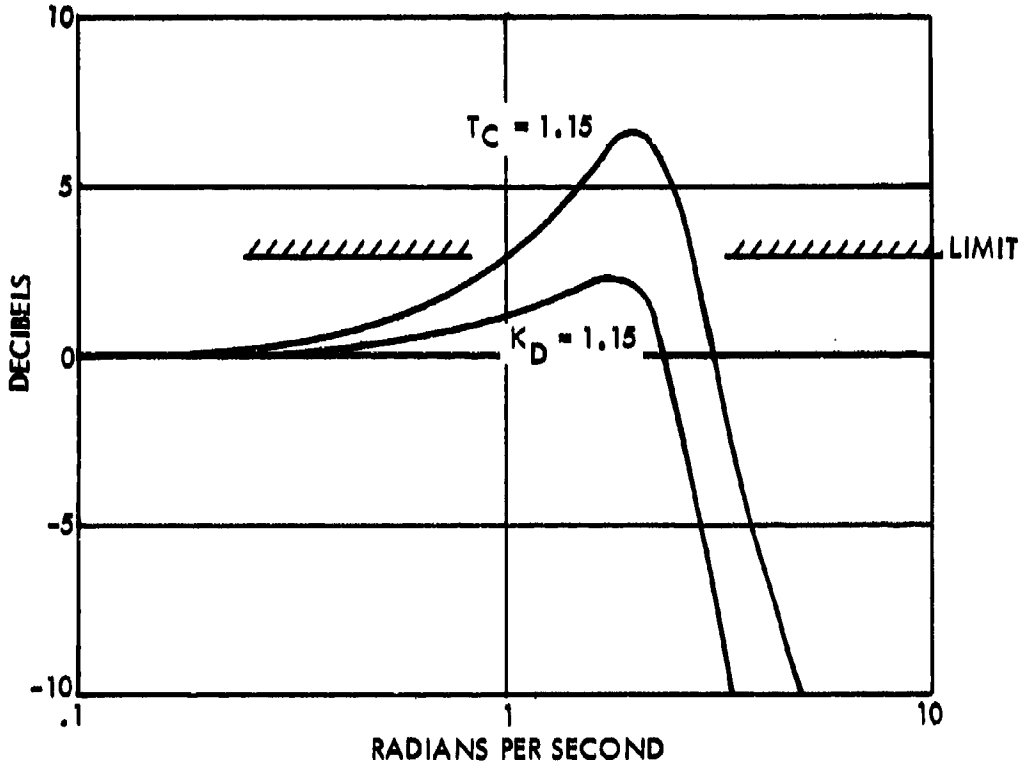


Figure 6. Closed-Loop Frequency Response

References

1. McRuer, D. T., et. al., A Systems Analysis View of Longitudinal Flying Qualities, WADD TR 60-43, January 1960.

2. Neal, T. P. and R. E. Smith, An In-Flight Investigation to Develop Control System Design Criteria for Fighter Airplanes, AFFDL-TR-70-74, Vol. I, December 1970.

3. Chen, R. T. N. and E. M. Boothe, An Investigation of Over-All Systems Criteria for the Longitudinal Flying Qualities of Highly Augmented Fighter Aircraft, AIAA Paper 74-833, August 1974.

4. Chalk, C. R., et. al., Revisions to MIL-F-8785B (ASG) Proposed by Cornell Aeronautical Laboratory, Inc., under Contract F33615-72-C-1254, AFFDL-TR-72-41, April 1972.

INVESTIGATION OF A CONTROL DEVICE WITH FORCE INPUT AND DISPLACEMENT FEEDBACK IN STEERING A LOW-FREQUENCY SYSTEM

Hans E. Boller, Walther Krüger
Forschungsinstitut für Anthropotechnik (FAT)
D-5307 Wachtberg-Werthhoven, F.R. Germany

Abstract

In manual control information about the vehicle dynamic behavior can be advantageous for the operator. With low frequency systems it must be determined wether the advantage of kinesthetic feedback can be utilized for the slow movements, because technical complexity and costs involved are at a relatively high level. This studie compares two two-handed yoke controls with the control functions of (1) displacement input and (2) force input with positional feedback as a kinesthetic display of vehicle response. In experiments, eight subjects had to reach command values as quickly as possible controlling the low frequency system output variable and its derivative (depth and vertical rate of a simulated submarine). The values obtained show that the technical complexity involved in a control with force input and positional feedback is not worthwile for the single-axis vehicle dynamic chosen.

Introduction

In man-machine-systems the main task of the human operator is often manual control. Even if vehicle steering is done automatically, in certain situations the human operator has to take over totally or partly (semi-automatic). Design of displays and controls should be compatible with the psychophysical characteristics of the human controller.

Figure 1 shows the principle relationships of a man-machine system. The operator receives display information about the desired and actual states of the system, and by means of control devices he makes inputs to the system as required by the differences between displayed values and dynamic plant characteristics.

In every human control action there is also proprioceptive feedback. This nonvisual feedback signal (upper dashed path) provides passive information about the state of the control, characterized by its movement resistance [1], [2].

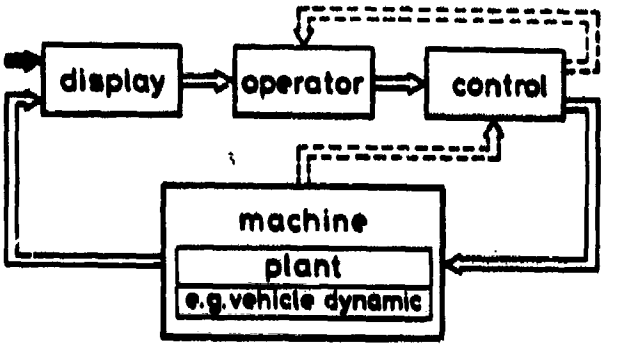


Figure 1 : Blockdiagram of a simplified man-machine-system

In addition through proper design the control device can provide feedback (lower dashed path) from the controlled system by making the control active. In the active case operators force is the input signal, while position (displacement) of the control device is driven by the system [3] , [4] .

With use of such a technique the operator is given a positioning task by overcoming forces of various strengths in the case of fast reacting systems. For low frequency systems the "positioning feeling" is not evident because of the long lag times. However, kinesthetic information remains in addition to the visual channel.

In order to evaluate the advantages of a force control device with position feedback in a low frequency system the following investigation was conducted simulating a submarine type vehicle.

Submarine control is a very complex task for the human operator. He has monitoring decision-making and controlling tasks. For this reason it is very important to look very carefully at the interface between the human and the machine during the conceptual phase. Different proposals have to be considered. Just so at the fixations of the function of the control device.

Posing the problem

Submarine depth is dynamically changed by displacement of bow and stern planes ; course by the rudder. The simulated submarine in our experiments had a semi-

automatic mode (H-mode). That is the submarine is controlled by varying vertical rate and turn rate. The operator controls these parameters by making inputs via a yoke to the so called "steering computer", which continuously determines from these inputs and instantaneous state variables the required motions of the rudder and planes.

Good results by MERHAV [4] with active kinesthetic feedback in high frequency systems led to our hypothesis that similar benefits may possible with low frequency systems. With kinesthetic feedback of vertical rate and/or course as conveyed by the active motion of the control yoke the visual channel of the operator might be unburdened. This could result in better control performance. But because of additional control complexity and costs, it is necessary to evaluate cost-effectiveness of alternative control methods before developing a new control device.

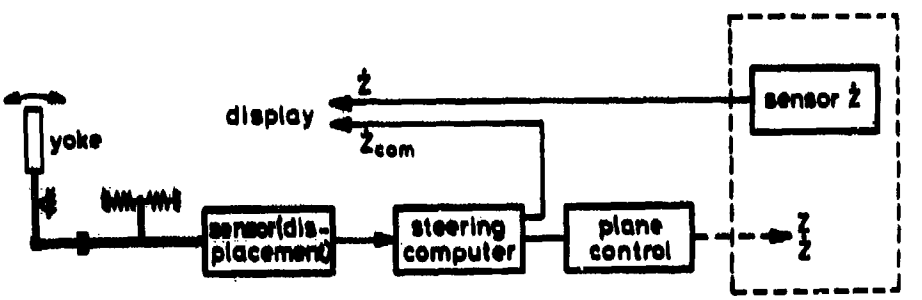


Figure 2 : Depth control by a spring centered control (passive)

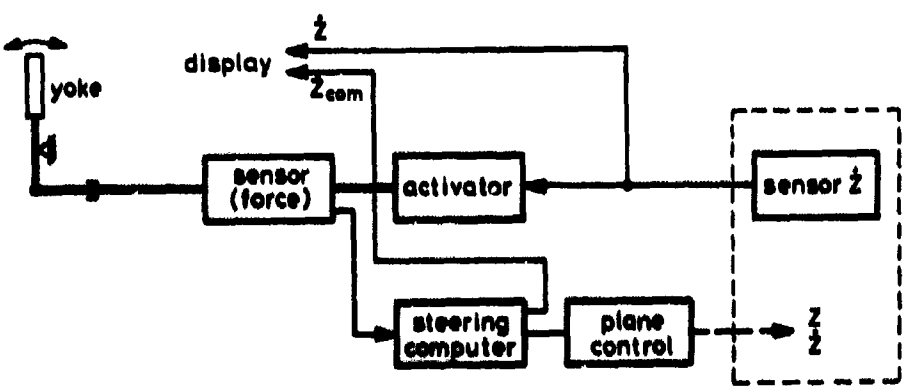


Figure 3 : Depth control by force input with displacement feedback (active)

In figures 2 and 3 schematic outlines of our alternative control devices are presented. Figure 2 illustrates the control device without active kinesthetic feedback. The displacement values of the yoke are measured by a sensor and transmitted to the "steering computer". This control device is called passive because the kinesthetic feedback is only passive information about the control device itself. Figure 3 shows the control device with active kinesthetic feedback. The input to the yoke is now measured by a force sensor. The information is transmitted to the "steering computer" as with the passive control device. But now there is a feedback of vertical rate or turn rate to an activator, which is mechanically moving the yoke proportional to these values. We have called this control device with active kinesthetic feedback the "active" control device.

A comparison of these two yokes in their control function and performance was made with an analog, linear simulation of a submarine.

Controlled plant

The basis for the simulation of the submarine dynamic are the equations of GERTLER [5] . It was necessary to only consider the depth axis, because that is the more difficult one. At 10 kts. the submarine will be controlled in depth only by the sternplanes. The transfer-function for our system which is a very low-frequency system with great instability is :

$$G(s) = \frac{Z(s)}{\delta_s(s)} = \frac{b_o + b_1 s + b_2 s^2}{a_1 s + a_2 s^2 + a_3 s^3 + a_4 s^4}$$

$Z(s)$ = depth = output

$\delta_s(s)$ = sternplane displacement = input

Figure 4 represents the maneuver for a "vertical overshoot" at 10 kts. with a sternplane angle of $\pm 12^{\circ}$.

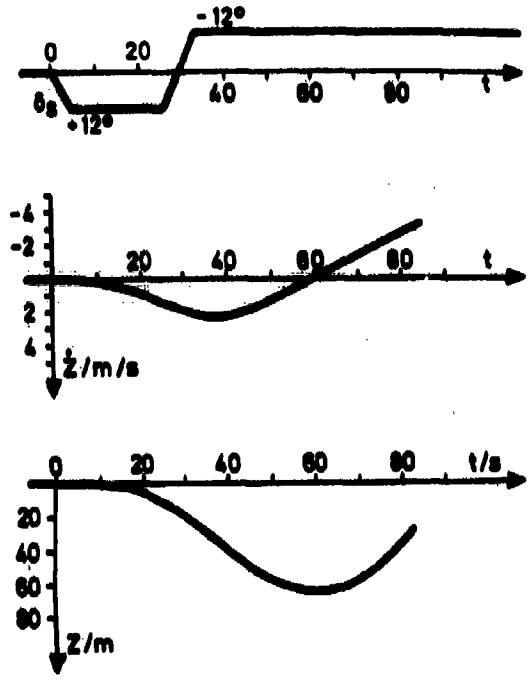


Figure 4 : Maneuver "vertiale overshoot" at 10 kts.

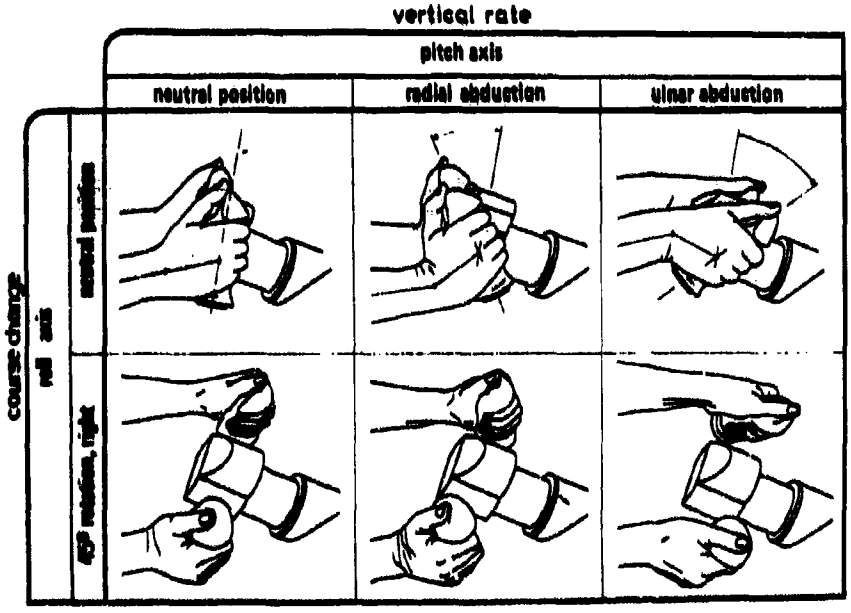


Figure 5 : Control yoke (from GÄRTNER [6])

Experimental set-up

The control yokes (see fig. 5) have two rotating axes : course and depth, but only the depth axis is used in this experiment. The yoke is linked to a torque-motor, which is controlled by an analog computer. This equipment makes it possible to reproduce either an isometric control device with displacement feed-back or a spring-centered device. The active device is damped electrically by a rate controller and in addition by a mechanical damping. On both control input force is measured with strain gauges. After exceeding a level of ± 1.78 Nm (dead zone) an integrator is used to drive the command pointer for vertical rate $\dot{Z}$ (see fig. 6). The spring constant of the spring-centered control device is c = .65 Nm/degree.

If there is a force or displacement on the yoke, which is higher than the level of the dead zone, the command value of $\dot{Z}$ will be changed proportional to the input force or displacement. The PID-controller compares command and actual values and generates control signals with less than a 5 percent overshoot of vertical rate even with step command inputs. The running time for the planes for full deflection is 12.5 seconds (± 25 degrees).

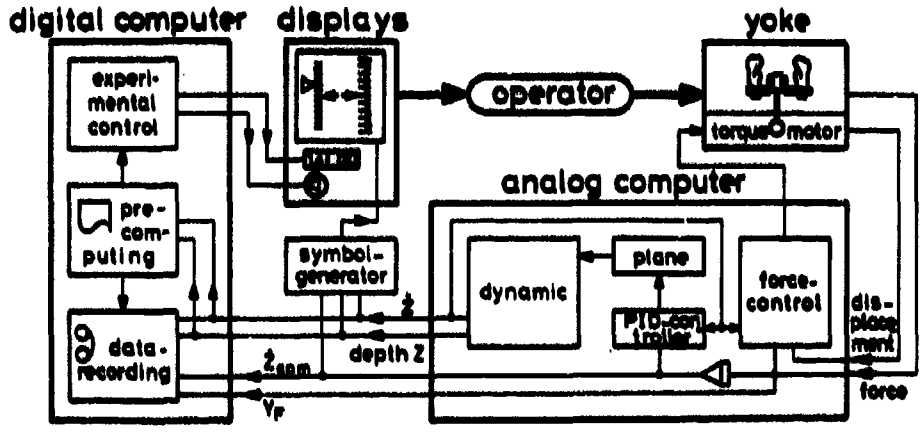


Figure 6 : Block diagram of the experimental set-up

Figure 6 shows the connections of the experimental set-up. The simulation of the submarine, the controller for the vertical rate, and the control of the

torque-motor are done on the analog computer. The outputs of the analog computer are the state variables Z and $\dot{Z}$ which are displayed and recorded. Ordered depth commands to the subjects are generated by a digital program and displayed numerically. In addition, the changes of ordered depth are acoustically indicated by a buzzer.

Experimental Method

The intention of the semi-automatic mode is to provide operator unburdening without restricting the use of the full dynamic variability of the submarine. The main operator task is to change depth and especially to do this as quickly as possible. Accuracy in reaching the new depth is not so important in many mission situations. But for measurement purposes the operator must be motivated to also stabalize as quickly as possible at the new depth and a depth "window" must be established as a measurement criterion for saying when the depth change maneuver is completed. For these reasons we changed the task description for the subjects accordingly.

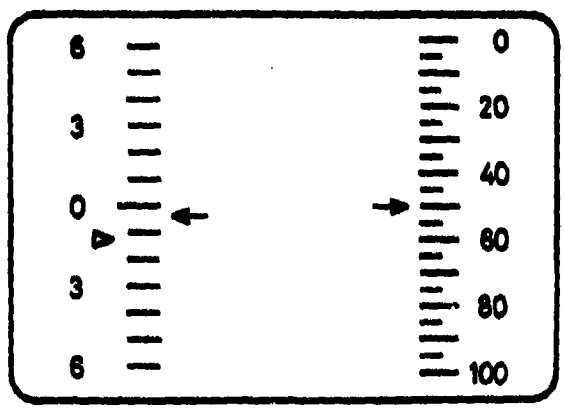


Figure 7 : Used display with the scales of vertical rate and depth

The subject's display (fig. 7) incorporates a depth scale and a vertical rate scale with the command and actual rate pointers. Time histories for a depth transition are shown in figure 8. After the order to change to a new depth the subject produces a force at the yoke. As a result the command pointer of the vertical rate changes, but only if yoke force exceeds a minimum level. The

simulated boat is next to respond and the subject has to judge when it is appropriate to "pull-out" for the new depth. When the subject stabalizes at the new depth within a pre-determined tolerance range for a fixed time, the maneuver is considered completed. These criteria values are not known by the subject. A further requirement during this maneuver is to nullify $\dot{Z}_{actual}$.

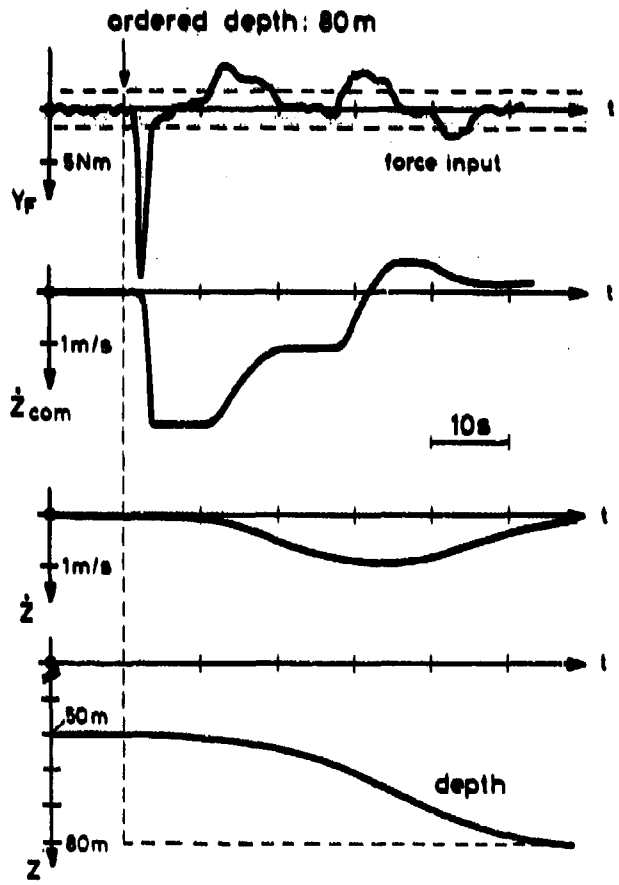


Figure 8 : Schematic time histories during a depth-transition

The time parameters and tolerance ranges are defined in figure 9. The ordered depth is considered to be reached, when the depth window expression $|Z - Z_{ordered}| < 1.2$ m during 8 s is fulfilled. T_o is the time required to reach a distance of 5 m from the ordered depth, which we call the approximation time. The first entrance into the Z-window is measured by T_1, the definite entrance is measured by T_4. In example A, $T_1 = T_4$, in example B, T_4 is

greater than T_1. The time after T_0 until the end of the Z-window is called T_{REF}. During this time root-mean-square (RMS) depth error will be computed (adjustment error). This is the effective depth error occurring between the time of reaching the 5 m distance until the end of the Z-window.

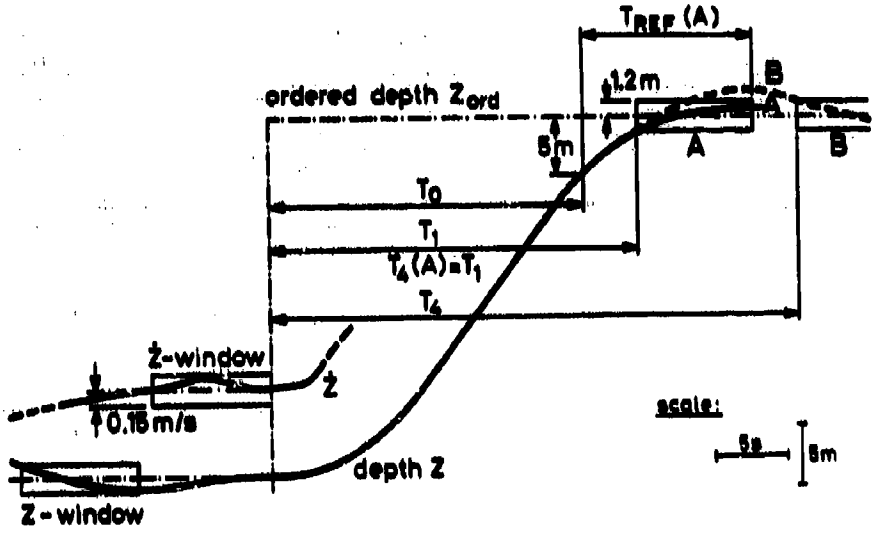


Figure 2 : Definitions for a depth transition

In addition to the main task a side task was used to enhance our measurement of performance and testing of the hypothesis that there is visual unburdening by active proprioceptive feedback. The "warning light" side task of EPHRATH [7] was selected. It consists of two small red lights mounted above each other outside the subject's peripherical visual field. During the run the upper or lower light, with equal probability, lights up and is followed again by another illumination after a duration varying randomly between .5 and 5 s. A correct response by the subject is to turn off the light by pushing the appropriate push button mounted on the left side near the yoke. The turn-off response must be given within two seconds of illumination. If not, the light extinguishes. Reaction time and number of wrong responses were measured and workload index was calculated.

To counterbalance transfer effects we used two groups of four subjects each and each group started with a different type of control device. The use of four

subjects was judged to be enough for determining the practical performance differences (rather than statistical significances) to help us to answer the question of the cost-effectiveness of the control device with active kinesthetic feedback.

After training, each subject had to perform 6 runs with 6 depth transitions in each run for each type of control device. Finally, there were two additional runs for each subject with both control devices to permit the collection of comparative subjective data, that is subjective ratings and questionnaires.

Each experimental run lasted about 10 minutes. Two runs were made daily in the mornings. A pause of 5 minutes was given between runs.

Results

To measure how quickly ordered depth changes are made the time parameters T_0 and T_4 are the most important dependent variables. In figure 10 the average approximation times T_0 and the average adjustment times T_4 are shown for each of the runs for both groups. Each data point represents the average of 4 subjects. Remarkable differences between the test conditions force-control and displacement-control were found out neither in the approximation time T_0 nor in the adjustment time T_4.

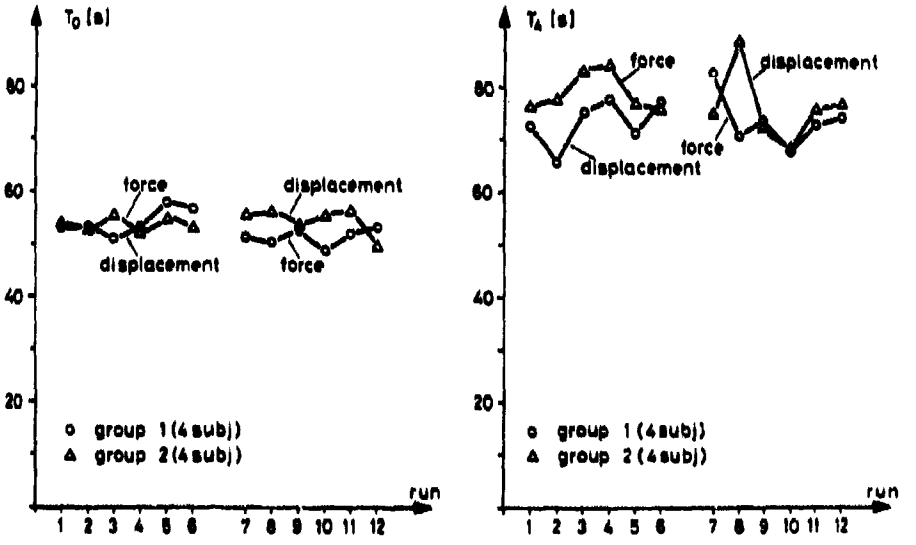


Figure 10 : Approximation time T_0 and adjustment time T_4

Another criterion for the property of a depth transition may be given by the adjustment error REFE (RMS depth error during T_{REF}). Figure 11 shows that there are no significant differences between the two test conditions.

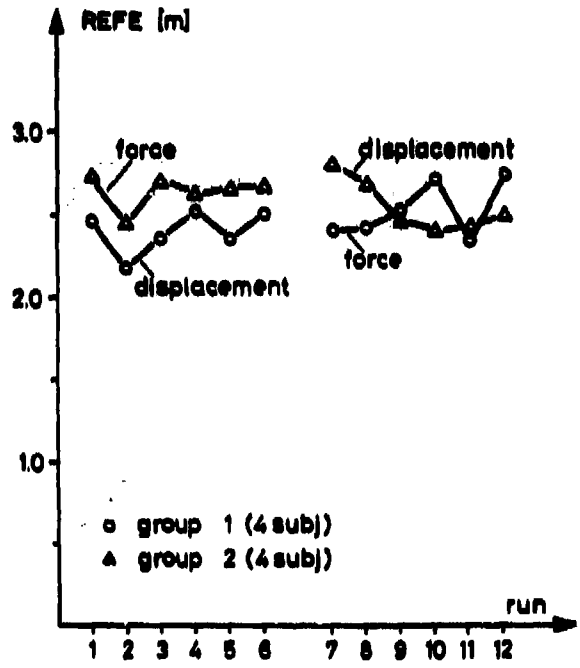


Figure 11 : Adjustment error REFE

The next step was to evaluate the side task to get a workload index for the main task [8] . This workload index consists of two parts [7] : (1) RTR, the response-time ratio for both "hits" and "misses"

$$RTR = \frac{\text{cummulative latency } (\Sigma_i\ T_i)}{\text{total number of stimuli} * 2\ s}$$

and (2) MR, the miss-rate

$$MR = \frac{\text{number of stimuli missed}^{+}}{\text{total number of stimuli}}$$

After weighting the workload index is

$$WLX\ [\%] = 55.5\ [\%] \cdot RTR + 45.5\ [\%] \cdot MR$$

+ incorrect and missed

A workload index of 100 % means that the subject only performes the main task.

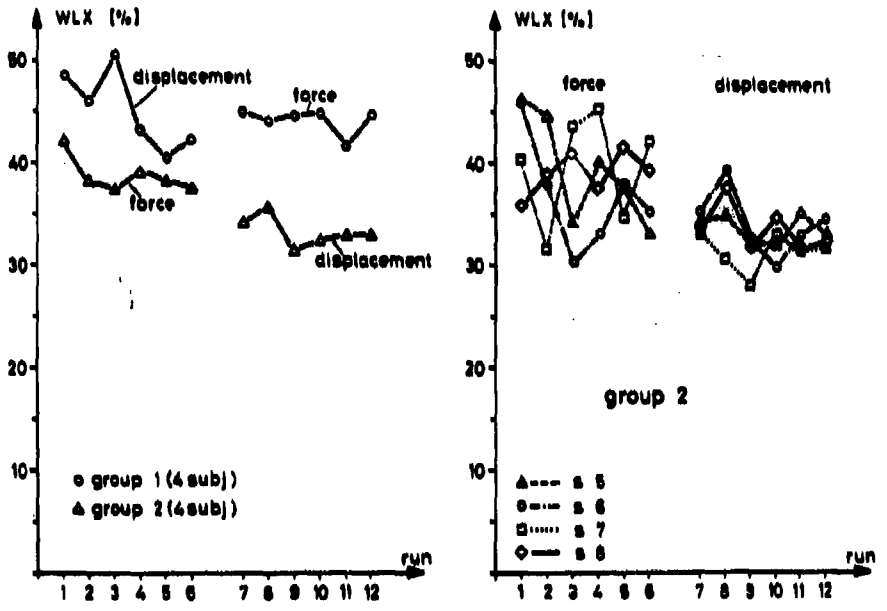


Figure 12 : workload index WLX

Figure 12 shows the workload index dependent on the single runs for each group. On the average group 2 has a difference between the two conditions force input (WLX = 39.7 %) and displacement input (WLX = 33.3 %). The workload index representation for each subject of group 2 let suppose that the difference is not established by the yoke functions, but by random variations. The verification by the WILCOXON-Test proves that the difference is not significant.

In the subjective rating the subjects had to judge the difficulty of the task and the difficulty of command pointer adjustment dependent on the function of the yoke. In addition they had to judge how much each control device was helpful or embarrassing. The results are shown in figure 13. The performance of the task with the "force input" control device was judged to be a little more difficult (a). In moving the command pointer there was no priority given by the subjects (b), but the active movement of the yoke was ranked as a little helpful (c). In an additional questionnaire the subjects had to give a written commentary and to circle answers of questions. Hereby the tendencies of the ratings were supported.

judgements: 1 ≙ a little, 2 ≙ some, 3 ≙ very much

	a			b			c		
	The task with the force input was			The adjusment of the cursor with the force input was			The displacement feedback was		
Vp	easier	0	harder	easier	0	harder	help	0	hindrance
1			1		0				2
2		0			0			0	
3			1		0		-1		
4	-1			-1			-1		
5			1			2		0	
6			1	-2			-2		
7			3			3			3
8	-1			-1			-2		
	median: 1			median: 0			median: -0.5		

Figure 13 : Subjective rating

Conclusions

This study compares two different single-axis control yokes, each with the functional capability of either pure displacement input or force input with a displacement feedback as a kinesthetic display of vehicle response. Because of relatively high technical complexity and costs of the displacement feedback, the goal of the investigation is to determine practical performance differences. In experiments subjects had to reach command values as quickly as possible without a large overshoot controlling the output variable of a low-frequency system and its derivative, which was also displayed kinesthetically in the displacement feedback condition.

Results are based on transition times, adjustment errors and indices of a side task. Statistical analyses and subjective ratings and questionnaires did not support a preference for the force yoke with a displacement feedback to such an extent as it would be necessary to justify the relatively high technical costs. Concerning both, human operator performance and costs, it is therefore

recommended to use a control yoke with pure displacement input in steering one-axis low frequency systems like the one investigated here.

References

[1] KRAISS, K.-F. : Can proprioceptive cues unload the human operator ? 6th Annual Conference on Manual Control, Proc., Wright Patterson AFB, Ohio, 7-9 April 1970, S. 535-547

[2] ROTHBAUER, G. : Influences of joystick spring resistance on the execution of simple and complex positioning movements. 13th Annual Conference on Manual Control, Proc., MIT Cambridge, Mass., June 15-17, 1977

[3] HERZOG, J.H. : Manual Control Using the Matched Manipulator Control Technique. IEEE Transactions on Man-Machine Systems, Vol. MMS-9, No. 3, Sept. 1968, S. 56-60

[4] MERHAV, S.J. ; YA'ACOV, O.B. : Control Augmentation and Work Load Reduction by Kinesthetic Information from the Manipulator. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-6, No. 12, Dez. 1976, S. 825-835

[5] GERTLER, M. ; HAGEN, G.R. : Standard Equations of Motion for Submarine Simulation. Naval Ship Research and Development Center, Report 2510, June 1967

[6] GÄRTNER, K.-P. ; RAU, G. : Biomechanical Analysis Applied in Designing the Operating Range of Hand Controls. In : E. Asmussen and K. Jørgensen (Eds.), Biomechanics VI - B, pp. 217-223, University Park Press, Baltimore, 1978

[7] EPHRATH, A.R. ; CURRY, R.E. : Detection by Pilots of System Failures During Instrument Landings. IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-7, No. 12, Dec. 1977

[8] ROLFE, J.M. : The Secondary Task as a Measure of Mental Load. In : W.T. Singleton et al. (Eds.), Measurement of Man at Work, London, Taylor and Francis, pp. 135-148, 1971

OPTIMIZING SPRING FORCE, VISCOUS DAMPING, AND INERTIA OF A JOYSTICK

Walther Krüger
Forschungsinstitut für Anthropotechnik (FAT)
D-5307 Wachtberg-Werthhoven, F.R. Germany

Abstract

In manual control, the performance of the human operator can be influenced by the design of the control device. Displacement-resistance characteristics are of special interest due to the movement resistance of the control. In order to investigate several reaction forces of a joystick, a series of experiments was conducted. Thereby, different experimental methods were used and compared. One goal of the experiments was to find optimal control forces as a combination of spring force, viscous damping, and inertia in continuous tracking of second- and third-order systems. Results show that only spring force (gradient: 1. - 3.5 N/cm) can improve tracking performance by 30 percent. Combining the low spring stiffness of 1. N/cm with damping (.35 N/cm) and/or inertia (2. kg), tracking performance was reduced. Using higher stiffness, no influence of damping and inertia could be infered. During one experimental condition, subjects had the possibility to adjust spring stiffness. Well trained subjects were choosing a value of 1.7 N/cm, on the average.

Introduction

In manual control proprioceptive feedback should be considered in designing control resistance of the stick. Before the development of servo-systems, a natural control feeling resulted because of the rigid mechanical link of the control to the vehicle's control surface (e.g. aircraft elevator). Current servo-systems eliminate this control feel thereby requiring simulated feel forces through artificial feel systems.

The occurrence of proper movement resistance would contribute to control feel and to operator performance. Warning information and quantitative feedback of system state can also be transduced by proper design of the control device in order to unburden the human operator's visual tasks.

The movement resistance of the control

When the operator executes a control movement, his input force F is opposite to the movement resistance R of the control. There are four different types of resistances [1] caused by

- friction C
- spring constant K
- viscous damping D
- inertia M

Total movement resistance R is expressed by the relationship:

$$R = C + K \cdot r(t) + D \cdot \dot{r}(t) + M \cdot \ddot{r}(t),$$

whereby, r(t) means the deflection of the control, $\dot{r}(t)$ the rate, and $\ddot{r}(t)$ the acceleration.

In general, every control has more than one type of control resistance, because all controls have some mass and they are rarely frictionless. F r i c t i o n usually is disadvantageous in manual control, especially for precise adjustments [2] and should therefore be minimized. Since friction is only desired in unusual cases, its effects are not considered in the following investigation. With s p r i n g s, which have a restoring force proportional to the deflection, the operator can easily find the zero-position of the control when he doesn't want to make control inputs to the system. The higher the spring constant (i.e. stiffness of the spring), the better is the centering force of the joystick. With viscous d a m p i n g and i n e r t i a, the moving resistance of the control becomes proportional to rate and acceleration of the movement. But it is questionable whether the operator can use this information precisely.

Although the importance of movement resistance in manual control has long been known, especially in aircraft control [3], the few investigations, which have been made on this topic, only consider some of the aspects [4], [5], [6], mostly using relatively simple control systems [7], [8], [9], [10], [11].

Evaluating procedures

To evaluate the effects of movement resistance on operator performance in a practical tracking task several measurement procedures were used and compared. Thereby, a range of fixed values for the three resistance parameters were selected and the parameter combinations were presented to the subjects. The performance measurement was done in three different ways :

(1) The first measurement method : experiments with fixed task difficulty. This standard method is realized by measuring tracking error.

(2) The second measurement method : experiments with fixed task difficulty and side task. The subject has to perform the main task as well as he can and he tries to simultaneously perform the side task when he can. The possibility of measuring performance with this approach is based on the limited signal processing capacity of the humen operator. Thus, for example if no performance differences appear in the main task, side task measures provide workload information in the main task [12]. The lower the performance in the side task the less is the signal processing capacity remaining to the subject. This means, that more effort was necessary in solving the main task.

(3) The third measurement method : experiments with difficulty varying as a function of performance (adaptive simulation [13], [14]). The subject is forced to operate near his performance limit, i.e. he is almost totally loaded and acts nearly without remaining capacity. Using the adaptive technique, average performance during a certain time is measured in order to change task difficulty : relatively low tracking errors lead to higher difficulty levels, large tracking errors reduce task difficulty. Thus, the obtained difficulty can serve as the dependent variable.

Another way to get stick resistance values could be reached by the following : Instead of presenting fixed independent variables a final approach permitted the subject operator to adjust the movement resistance parameters himself. Consequently subjects should know the whole parameter range. To avoid biased adjustments subjects should adjust the parameters starting from both high values as well as low values so that the mean of the different values obtained can be used [15].

Tracking task and experimental set-up

The control task used in the experiments is similar to aircraft control during landing with ILS in turbulent air. Dynamic characteristics of the aircraft are much simplified and control axes are uncoupled. Thus, the plant is third order in the lateral and second order in the vertical axis.

To avoid stabilization problems, an artificial horizon was displayed, which was a first order system with respect to the joystick deflection. An aircraft symbol was used as the reference mark. The deviation from the commanded path was shown in the form of a crosspointer.

The operator control station is connected with a hybrid computing system via an analog data transmission line (see figure 1). The analog computer (EAI 680 contains the simulation of the vehicle, the program to generate the display signals, and the computation of the control force. All other experimental control aspects including error measurement and parameter variation is done by the digital computer (SIEMENS 306).

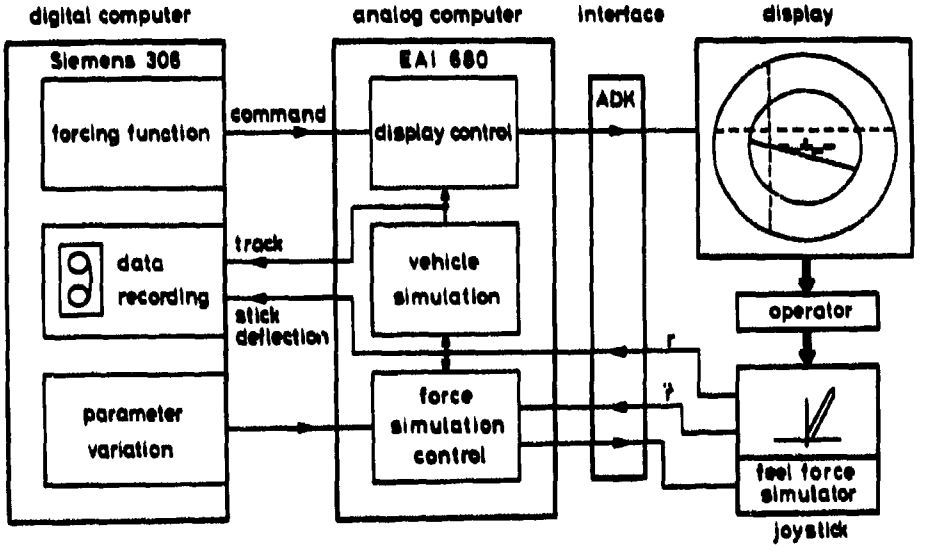


Figure 1 : Block diagram of the experimental set-up

Figure 2 is a photo of the subjects control station. The control used is a joystick of 70 cm length, that can be deflected 15 cm maximal in each direction from its center (neutral position). The movement resistance is generated by an electrohydraulic system. The distance between joystick and back of the pilot-seat is 55 cm. Subject's eyes are 120 cm away from the display screen, which is a television monitor behind a circular section of 34 cm diameter.



Figure 2 : Experimental arrangement

Experimental procedure

Various associates were used as subjects in the experiments. Two subjects were used with the performance measurements presenting fixed resistance parameters. Five subjects were employed with self-adjusted resistance parameters. The low number of subjects results from the practical fact, that in very difficult task well trained subjects provide better data than relatively untrained subjects. Training time is extremely large in relation to the total experimental data collection time. Reasonable statistical inferences are still possible by using a large number of runs for each subject.

The subjects' task was to minimize the deviation between cross pointer and reference aircraft symbol (compensatory tracking task). During training they only had to compensate self-induced deviations. During experimental runs, the task was more difficult because of the addition of a forcing function (simulation of simplified gusts). The same stochastic forcing function pattern was used in all tracking runs. When four consecutive runs of a certain movement resistance showed nearly no learning effects, they were taken as the score for this parameter combination. Analysis of variance was applied to the data collected.

Results with fixed task difficulty

The following results are from the method using fixed parameters of the control resistance and fixed task difficulty. The movement-resistance consists of a parameter combination from the independent variables

spring constant (0, 1., 2., 3.5 N/cm),

viscous damping (0, .2, .35 N/cm/s),

inertia (0, 2. kg)

The dependent variable was root mean square (RMS) tracking error.

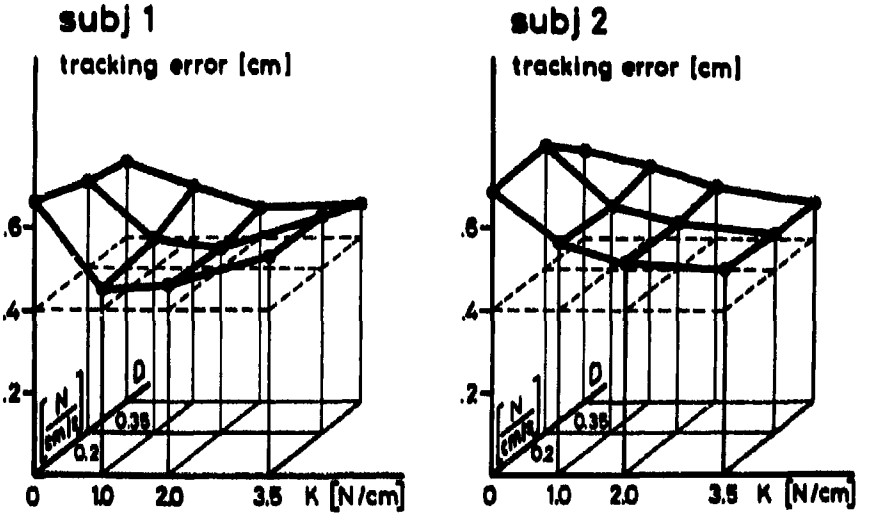


Figure 3 : Tracking error as a function of spring constant K and viscous damping D

Figure 3 shows the tracking error depending on spring constant K and viscous damping D of the joystick. In the first experimental block, there was no joystick inertia. Each point in the diagram represents the mean of four two-minutes runs.

It can be seen, that tracking error is significantly greatest with no joystick centring force, i.e. zero spring constant. Viscous damping of the joystick had no consistent influence on tracking error, but there was a tendency to greater error with spring constant 1. N/cm. There was a strongly reduced tracking error with subject one using spring constants 1. to 2. N/cm with low or zero damping. For subject two, this reduction was less.There was no significant changing of the tracking error for spring constants greater than 2. N/cm, regardless the various damping values.

In the next block of experimental runs, the influence of spring constant and viscous damping was investigated using a joystick with 2 kg inertia.

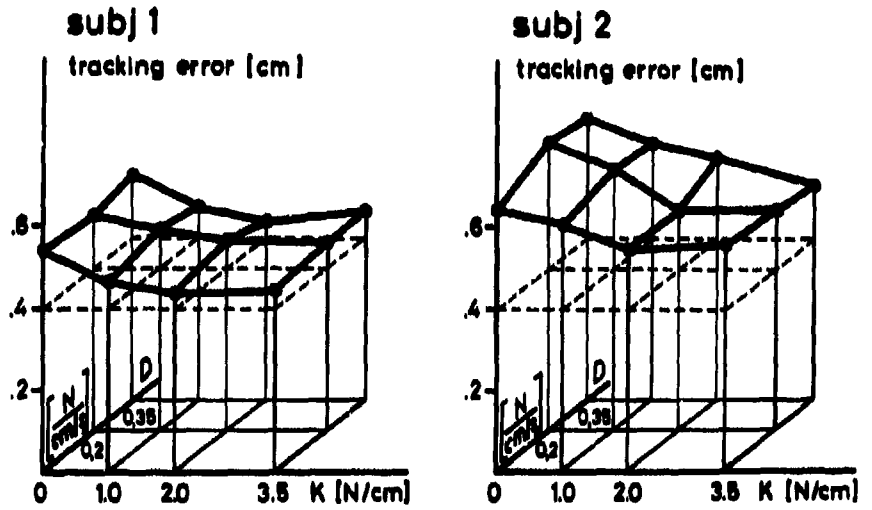


Figure 4: Tracking error as a function of spring constant K and viscous damping D with inertia (2 kg)

Principally the results (see figure 4) are the same as in the runs without inertia, but performance differences were even smaller.

Another presentation of the results is given in figure 5. Damping data were pooled and tracking error as a function of spring constant is shown for both subjects individually (dashed lines) and averaged (complete lines).

On the average, the significantly decreasing of tracking error was attained as the spring constant increased to 1. N/cm for the zero inertia condition and was attained as spring constant increased to 2. N/cm for the "with inertia" condition. Tracking error remained the same for all the three spring constants for the zero inertia condition, but for the 2 kg inertia condition only the spring constants 2. and 3.5 N/cm showed the same low tracking error. The decrease of tracking error from zero centering force to spring constant 2. N/cm amounts 32 percent without inertia and 21 percent with inertia.

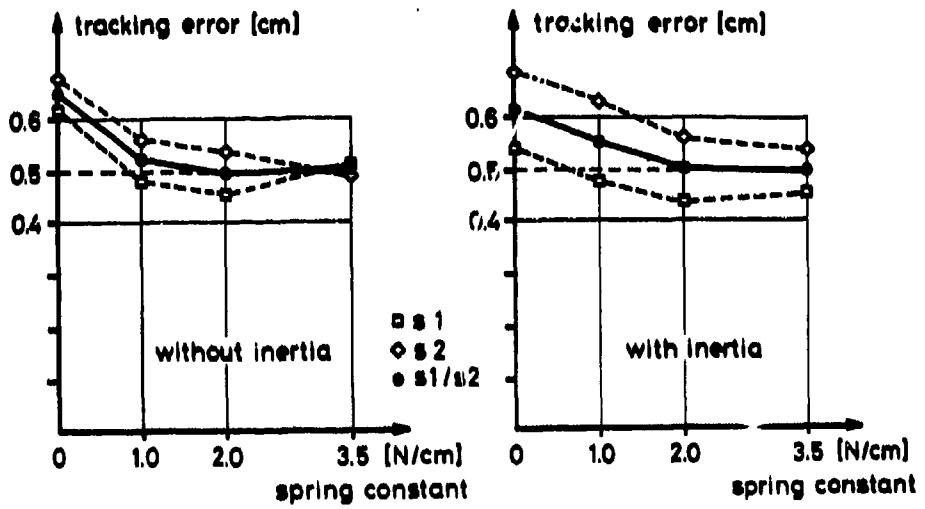


Figure 5 : Tracking error as a function of spring constant without and with inertia

Results with other measurement procedures

In order to assure proper measurement while investigating the effect of joystick movement resistance on performance further experiments were conducted with other measurement techniques described earlier.

In one experimental condition subject's workload was increased by a side task. The attitude of all perpendicular display components was angled from time to time, and the subject had to bring the attitude back to perpendicular as fast as possible by turning a knob. The correct perpendicular was indicated by the extinguishing of a small light. The total time that display components were off the perpendicular remained nearly constant for all parameter combinations tested. The main task tracking errors between the various parameters had the same relationship as during the experiments without side-task.

Another experimental condition was conducted using the adaptive technique. The difficulty of the tracking task, caused by the cut-off frequency of the forcing function, varied depending on the RMS error of the human operator. The task difficulties which subject operators were able to handle were larger with higher spring constants in the same relationship as the error results of the previous experiments.

During self-adjustment experiments subjects adjusted spring stiffness by rotating a knob mounted on the arm rest of the pilot's seat. In these runs, damping and inertia were not varied because of the relatively unimportant influence of these parameters. Five subjects participated. Three of them were well trained, and they consequently selected on average a spring constant of 1.7 N/cm (.22 N/cm standard deviation, based on 30 runs). This value falls within the range of 1. to 3.5 N/cm which was found to have lower errors in the previous experiments. Considering that subjects had to approach the desirable value starting from both a upper value of 5.N/cm and a lower value of zero the 12 percent standard deviation is rather low.

The results of the less trained subjects indicated that adjusted values were biased more by initial values. Thus using the lower value as starting position, the mean adjusted value was 1. N/cm, but coming from the upper value, 1.4 N/cm resulted.

Conclusions

The objective of the experiments was to measure performance as a function of movement resistance parameters and to compare the results of several measurement techniques.

In general it was seen that the various movement resistance parameters of the joystick that were tested had no significant influence on tracking performance with one exception. Changing over from an unrestrained joystick to one with a defined deflection resistance, especially, with a spring constant of more than 1. N/cm without inertia or 2. N/cm with inertia led to a clear and significant performance increase. Combining of the low spring stiffness of 1. N/cm with either damping (.35 N/cm/s) or inertia (2 kg) reduced tracking performance. With spring constants above these minimums no performance improvement or decrement was found. Damping and inertia also had no influence on performance with the higher spring constants.

In conclusion, the results indicate that the control device should have some minimum centering force, but its intensity is not critical in a wide range beyond this value. If inertia or damping exists in a control, minimum centering force should be stronger.

Other measurement method results support these same conclusions obtained with the standard measurements. Both, the additional loading of the subjects by a side task and the adaptive tracking showed identical effects of the experimental parameters.

Since tracking performance was relatively insensitive to the various movement resistance parameters including spring centering forces beyond a minimum value a low spring constant is recommended to minimize the physical loading of the operator. This recommendation is supported by the results of the finally used method, during which subjects self-adjusted spring stiffness to a mean value of 1.7 N/cm. The self-adjustment method is sufficiently exact and saves time and effort. It is therefore probably the most cost-effective approach to stick optimization.

References

[1] ELY, J.H. et al. — Design of Controls. Wright Air Development Center, WADC TR 56-172, 1956

[2] BROWN, B.P., REEDER, J.P. — Some Effects of Valve Friction and Stick Friction on Control Quality in a Helicopter with Hydraulic-Power Control Systems. Langley Aeronautical Laboratory, Langley Field, Va., NACA TN 4004, 1957

[3] ORLANSKY, J. — Psychological aspects of stick and rudder control in aircraft. Aeronaut. Engng. Rev., 1949, 8, 1 - 10

[4] WATLER, J.F. — Effects of static Friction and Force Displacement Gradient on a discrete visual Tracking Task. Long Beach, Calif., Douglas Aircraft Company Report LB 31050, 1963

[5] MAGDALENO, R.E., McRUER, D.R. — Effects of manipulator restraints on human operator performance. USAF Flight Dynamics Laboratory, Tech. Rept. TR-66-72, 1966

[6] BRIGGS, G.E., BAHRICK, H.P., FITTS, P.M. — The Influence of Force and Amplitude Cues on Learning and Performance in a Complex Tracking Task. AFPTRC-TN-57-33, Lackland Air Force Base, San Antonio, Texas, 1957

[7] WEISS, B. — The role of proprioceptive feedback in positioning responses. Journal of Experimental Psychology, Vol. 47, No. 3, P.215 - 224, 1955

[8] BAHRICK, H.P., BENNETT, W.F., FITTS, P.M. — Accuracy of Positioning as a Function of Spring Loading in a Control. Journal of Experimental Psychology, Vol. 49, No. 6, P. 437 - 444, 1955

[9] KRAISS, K.-F. — Can proprioceptive cues unload the human operator? 6th Annual Conference on Manual Control, Proc., Wright Patterson AFB, Ohio, 1970, p. 535 - 547

[10] HOWLAND, D., NOBLE, M.E.
The Effect of Physical Constants of a Control on Tracking Performance

[11] ROTHBAUER, G.
Influences of joystick spring resistance on the execution of simple and complex positioning movements

13th Annual Conference on Manual Control, Proc., MIT Cambridge, Mass., June 15-17, 1977

[12] ROLFE, J.M.
The Secondary Task as a Measure of Mental Load

In: W.T. SINGLETON et al. (Eds.): Measurement of Man at Work, London, Taylor and Francis, pp. 135 - 148, 1971

[13] KELLEY, C.R., KELLEY, E.J.
A Manual for Adaptive Techniques

Dunlap and Associates, Inc. Santa Monica, Calif., Aug. 1970

[14] STEIN, W., PIOCH, E.
Zur Anwendung von Regelaufgaben bei ergonomischen Untersuchungen

Forschungsbericht Nr. 23, Forschungsinstitut für Anthropotechnik, Meckenheim/Bonn, S. 93 - 108, 1975

[15] KRUSE, W., ROTHBAUER, G.
Optimization of Control Gain by operator adjustment

9th Annual Conference on Manual Control, Proc., MIT, Cambridge, Massachusetts, Mai 1973

DEVELOPMENT TESTING OF A MOVEABLE SIDESTICK CONTROLLER FOR THE PRODUCTION F-16

Terry L. Lutz, Captain, USAF
F-16 System Program Office (ASD/YPTT)
Wright-Patterson AFB, Ohio 45433

ABSTRACT

Since the nearly disastrous lateral oscillation that resulted in its first flight, the F-16 has undergone a series of changes to the unique sidestick controller that have sought the optimum blend of force and displacement. In May 1977, a controller was tested in YF-16 No. 2 that was a distinct improvement in the man-machine interface. This configuration was further refined and entered extensive testing in the F-16A to quantify the effect on flying qualities as a means of substantiating pilot comments. The result was a sidestick controller that reduced air-to-air tracking error by 40% while reducing pilot fatigue with slight motion and positive stop cues.

SESSION 5A: PILOT MODEL IDENTIFICATION

Moderator: Dr. William R. Wells, Wright State University

5a

ON THE IDENTIFICATION OF PARAMETERS IN THE OPTIMAL CONTROL MODEL

Roy E. Lancraft
Bolt Beranek and Newman Inc.
Cambridge, MA

and

David L. Kleinman
University of Connecticut
Storrs, CT

Presented at the

Fifteenth Annual Conference on Manual Control
Wright State University, Dayton, Ohio, March 20-22, 1979

Abstract

In order to validate (or invalidate) model assumptions and/or applications of the Optimal Control Model (OCM) it is necessary to match experimental data in an efficient and objective fashion. The ability to track individual parameter changes across different experimental situations is invaluable in understanding the model so that it may be used with confidence in new situations. This paper addresses these needs by examining the suitability and performance of using unconstrained non-derivative function minimization techniques to identify key human operator parameters in the Revised Optimal Control Model (ROCM). In particular, three methods will be explored. They include two of Powell's conjugate direction algorithims and a Quasi-Newton method. A sum of squares objective cost functional is adopted, and its properties discussed, along with convergence properties and the results of identification runs using each of the methods.

Introduction

"Tuning" parameters in the Revised Optimal Control Model (ROCM)* to match experimental data can be a tedious and somewhat subjective task if done manually. Therefore an automatic scheme to match data would be very desirable. Such a methodology is examined in this paper. The approach taken will be to identify the necessary parameters such that an objective cost function which embodies the available reduced order experimental data, is minimized. Unfortunately, the mapping between model outputs and model parameters is complex and non-linear. Since analytic derivatives are therefore not readily available, we chose to examine the

* The particular version of the OCM used in this study is a revised version in which control-rate is commanded, rather than control, and the "observations" contain direct control observations to represent proprioceptive cues to the controller. See [1,2,3].

performance of function minimization techniques which do not require derivatives. Those examined include, Powell's (1964) method for minimizing a general non-linear function, Powell's (1965) method for a sum of squares function, and a Quasi-Newton method.

The paper is organized into several sections. The first section presents the objective function to be minimized along with is properties. Next, the non-linear programming algorithms are briefly discussed. In the third section convergence properties of each of the methods are shown by example using model simulated K/s tracking data. Finally, the algorithms are applied to actual experimental data, and concluding remarks are made.

The Objective Function

Before stating the objective function we've chosen to minimize, let's summarize our parameter identification problem. An experimental "observation" is available. It consists typically of magnitude, phase, and remnant points at several frequencies, along with RMS system scores. The data obtained also consists of sample variances for each point. The sample variances result from averaging data across subjects and/or trials. Based on this "observation" we wish to identify a set of parameters, p, which result in model predictions to best match the observations. Further complicating the problem, the describing function and remnant points may be frequency correlated.

With the above considerations in mind, the following objective cost function (suggested by Levison [3]) is chosen to be:

$$J = \frac{1}{N_1}\sum_{i=1}^{N_1}\left[\left(\frac{G_i-\hat{G}_i}{SDG_i}\right)^2 + \left(\frac{P_i-\hat{P}_i}{SDP_i}\right)^2\right] WOX_i$$

$$+ \frac{1}{N_2}\sum_{i=1}^{N_2}\left(\frac{R_i-\hat{R}_i}{SDR_i}\right)^2 WTR_i + \frac{1}{N_3}\sum_{i=1}^{N_3}\left(\frac{S_i-\hat{S}_i}{SDS_i}\right)^2 WTS_i$$

where:

N_j = Number of non-zero weights in the j-th terms corresponding weighting vector.

G_i = Magnitude of the i-th describing function point to be matched.

P_i = Phase of the i-th describing function point to be matched.

R_i = Remnant of the i-th frequency point to be matched.

S_i = i-th RMS score to be matched.

"^" = "hatted" variables are the current estimates of those variables.

WOX, WTR, WTS = Weighting vector for describing function, remnant, and RMS scores terms respectively.

This choice of loss function gives a weighted least squares fit, where the weighting reflects confidence in the data. Minimizing this function is equivalent to minimizing the average number of standard deviations of the fitting error. Notice that each squared difference is normalized by the sample variance of the data, the number of points in the sum, and by a general weighting vector. The interpretation of each of these quantities, and their effect on the loss function, will now be discussed.

The sample variance represents the observed variability of each data point due to measurement errors and inter- and intra-subject variability. It also points out repeatable portions of the data. Dividing by the sample variance provides the required non-dimensionalization, allowing dissimilar terms to be summed together. It also automatically gives less importance to points with large variability.

Normalizing by the number of points contained in each sum has the effect of placing equal importance in matching the gain, phase, remnant and scores as groups of information. This was done for two reasons. First, the frequency domain points may be correlated, and as such each point doesn't represent a truly independent piece of information. Not knowing how to break up the frequency data into smaller collections of correlated points our first approach was to assume that all homogeneous data might be correlated. Secondly, each group is derived from the data using a different method, and equal weight (confidence) is given to each method.

One could argue that each of the rms scores are obtained by time averaging over the entire run, and therefore each measured score is equal in importance to each of the other groups. This was not examined, but does merit consideration. To account for the correlation of the frequency points, Wittenmark suggests that off diagonal terms might be employed in the weighting matrix, of the form $1 - e^{-\Delta/T}$. Where Δ is the difference in the frequencies, and T is a correlation time. Again, this was not tried but if a good estimate of the correlation time is available this approach may enable better interpretation of resulting parameter confidence intervals.

Inclusion of the weighting vector is due to practical programming considerations. The major purpose of the weighting vector is to allow bad data points (ones where the noise to signal ratio is very large) to be eliminated from the problem. This is accomplished by setting the corresponding weight (normally unity) to zero. However, the weighting vector is useful for another good reason. It allows the structure of the cost function to be changed at run-time, allowing investigation of any (diagonal) weighted least squares loss function. For example, consider using the weighting to cancel out the averaging. Then the loss function takes the form:

$$J = [z-h(p)]\ R^{-1}\ [z-h(p)]$$

where:

z = vector of observations
$h(p)$ = vector of model outputs
$R^{-1} = \text{diag}[\ 1/(SD_i)^2\ ]$

Minimization of this function would give the maximum likelihood estimate of the parameters p, if the observations were formed as

$$z = h(p) + V \quad ,$$

with, $V = N(0,R)$, noises independent, and R given by the sample variance of the data [5]. Note that in general the sample variances of the data will not equal the variance of the measurement noise because the measurement noise represents the innovations between the data (observations) and the model predictions.

In obtaining the loss function the objective has been to choose the weightings in a rational fashion, and in such a way as to approximate what we believe the actual innovations covariance to be. Since the observations are not independent and we cannot know the true innovations covariance, the estimated parameters will not be the minimum variance (maximum likelihood) estimates. However, if our engineering judjment is sound, the innovation covariance will be a good approximation, and our parameter estimates will approach the minimum varience estimates.

Non-linear Programming Methods

Now that the loss function has been determined we need to minimize it in an efficient and reliable way. This section briefly describes how the three function minimization techniques are performed, their theoretical properties and major computational requirements.

1. Powell's (1964) Method. [6]

Powell's method performs an unconstrained minimization of any non-linear function. Theoretically the method has the following properties:

- o If J is a general quadratic form, the minimum is found in a finite number of steps (approximately equal to the number of unknown parameters).
- o minimization is invariant under linear transformations of the parameter space.
- o Rate of convergence to the minimum is efficient in that if the original coordinate directions are poor, improved directions are found easily.

Basically, the minimization is carried out by finding pseudo-conjugate gradient search directions by perturbing each parameter individually (Euclidean basis space) at first, and in combinations later. At the end of each iteration a conjugate direction is formed which usually replaces an original direction. Theoretically, after n iterations (where n is the number of parameters) the directions are all mutually conjugate and a search along each direction will yield the desired minimum. Practically, however, an iteration may yield a dependent direction which should not become a replacement. Therefore, it may take more than n iterations to find the minimum. Unfortunately, a bivariate line search is needed to minimize J in each search direction because it is known only that a minimum lies along the particular conjugate direction. This means that for each iteration n bivariate line searches must be performed along each of the n directions.* Moreover, since the quadratic

* Depending on the efficiency of the line search (we used a modified DSC-Powell line search [8]) each iteration could require a large number of function calls.

form of the cost function is not recognized, this method doesn't yield any confidence limits.

2. Powell's (1965) Method. [7]

This method emulates the generalized least square method, with numerical gradients replacing analytic ones. Therefore, it is specific to sums of squares of non-linear functions. Some properties of this method are:

- o It has the same convergence properties as the generalized least squares method (quadratic) but doesn't require derivatives.
- o Only before the initial iteration are substantially more function evaluations required.
- o It is unlikely to converge before reaching the minimum.
- o Chooses conjugate directions of search.
- o It generates an approximation to the variance-covariance matrix.

Before the first iteration begins, n linearly independent directions (usually chosen as the Euclidean basis space) are required and the numerical gradients along these directions. After this initialization is done each iteration consists of the following computations:

- o Solving a linear equation to obtain a new search direction. (Note: if solved via partitioning this step requires N^2 rather than N^3 operations).
- o Searching along this direction to find the minimum, and at the same time computing an estimate of the derivative along this direction (Note: this requires no extra function evaluations).
- o Updating a single row and column of the hessian matrix.

Looking back, this means that a single univariate line search is required for each iteration rather than N bi-varitate line searches for each iteration of the 1964 Powell method.

3. Quasi-Newton Method

Simply stated the Quasi-Newton method is Newton's method using numerical derivatives instead of analytic ones. Therefore, the Quasi-Newton Method minimizes a function by searching at each iteration along the direction given by:

$$v = (\nabla^2 h(p))^{-1} \nabla h(p)$$

where the numerical gradients are formed as,

$$\nabla h(p) = \frac{h(p+\Delta) - h(p)}{\Delta} \equiv g(p)$$

and where the hessian is approximated by,

$$\nabla^2 h(p) \approx g(p)\, g(p)'$$

Rather than directly inverting the hessian matrix to form the next direction, an eigenvector decomposition is performed at each iteration enabling the inverse to be formed as

$$(\nabla^2 h(p))^{-1} = \sum_{i=1}^{n_p} 1/\lambda_i \; \xi_i \xi_i'$$

Where ξ_i is the eigenvector associated with the i-th eigenvalue, and where $n_p \leq n$. In forming the sum only those terms that correspond to eigenvalues no smaller than 10^{-5} times the largest one are included. This tends to eliminate some of the usual conditioning problems associated with eigenvalues spanning a wide dynamic range.* Note, if any terms are eliminated in forming the sum, the pseudo inverse results.

The algorithms converge if:

- o all parameter's relative change is less than some criterion level and,
- o the relative change in J is below a preset level or,
- o if J becomes very small or,
- o if a maximum number of iterations is exceeded.

Convergence Properties

To evaluate the convergence properties of each of the methods, identification runs were performed on a set of model generated K/s compensatory tracking data. Several different starting points in the parameter space were chosen to show efficiency and local identifiability. Also, in this way convergence to the true minimum could be monitored.

Table 1 presents the true model parameters used to generate the simulated tracking data. Notice that the observation noise ratios for error and error rate are not the same, as is the usual case. This was to simulate thresholds on these outputs. The other parameters are close to their usual nominal values. Since we are dealing with simple tracking data, we can assume (correctly) that mean squared error is minimized. Therefore, the key ROCM parameters were chosen to be the control-rate weighting, Qr (or equivalently T_N), time delay, T_D, motor noise $\rho_{\dot{u}}$, and observation noise variances, ρ_e, $\rho_{\dot{e}}$, ρ_u. It was felt that each of these parameters has a unique effect on model predictions; and if sufficient data were available, (i.e. enough high frequency phase to determine the time delay) all could be identified.**

---

* Another way to minimize the effects of (relatively) very small eigenvalues is to replace them by a constant. One choice could be to pick this constant as 10^{-5} times the largest eigenvalue.

** As a practical consideration Qr is identified along with the additive motor and observation noises. By doing so, iterations on T_N or noise-to-signal ratios need not be performed at each function call. Moreover, by invoking the

Because we are using model generated data, arbitrarily small standard deviations were uniformly applied to the data. They were chosen to be +0.5 dB on magnitude and remnant, +1.0 degree for phase, and +0.1% for the rms scores.

Table 1: Parameters used for Model Generated K/s Dynamics.

Conditions	Parameters						J
	$\rho\dot{u}$(dB)	ρs(dB)	$\rho\dot{e}$(dB)	ρu(dB)	T_N(sec)	T_D(sec)	
True parameters	-35	-18	-22	-25	.1	.2	
Initial guess #1	-25	-22	-22	-30	.1	.25	2700
#2	-40	-20	-20	-20	.08	.17	923
#3	-45	-22	-18	-30	.06	.15	3921

To begin our investigation, several initial parameter guesses were chosen in the neighborhood of the minimum so that convergence would occur from many different directions in parameter space. All of the initial guesses tried converged to the true minimum. This leads us to assume that the parameters are at least locally identifiable.*

Table 1 presents three typical initial guesses, along with their initial loss function values. All have large initial function values, and require many parameters to change before the minimum can be obtained. Table 2 shows the attained value of the loss function, confidence intervals, and identified parameter values for the three initial guesses. The model's match to the simulated data is shown in Figure 1.

Notice that all three methods, no matter what the initial guess, obtain the same parameter and loss function values. All the parameters are very close to their true values, with the exception of u. Since J is not very sensitive to u this is to be expected. This could also point out a bias in the parameters. However, since the simulated data had to be truncated when input to the identification program, this may account for the slight bias. Even though the same minimum is reached, the confidence intervals are different. Since the Quasi-Newton method

---

appropriate scaling assumptions the effective noise-to-signal ratios can still be reconstructed at the end of the identification. Depending on the order of the plant dynamics, the savings in computation time can be substantial.

* If we start too far from the minimum there is a possibility that the predictions of the phases (computed modulo 360 degrees) could be discontinuous in slope, or biased on the order of 360 degrees. The program attempts to maintain a smooth phase curve by keeping the predictions within 190° of the measured data points. But, far from the minimum this may not be a good assumption.

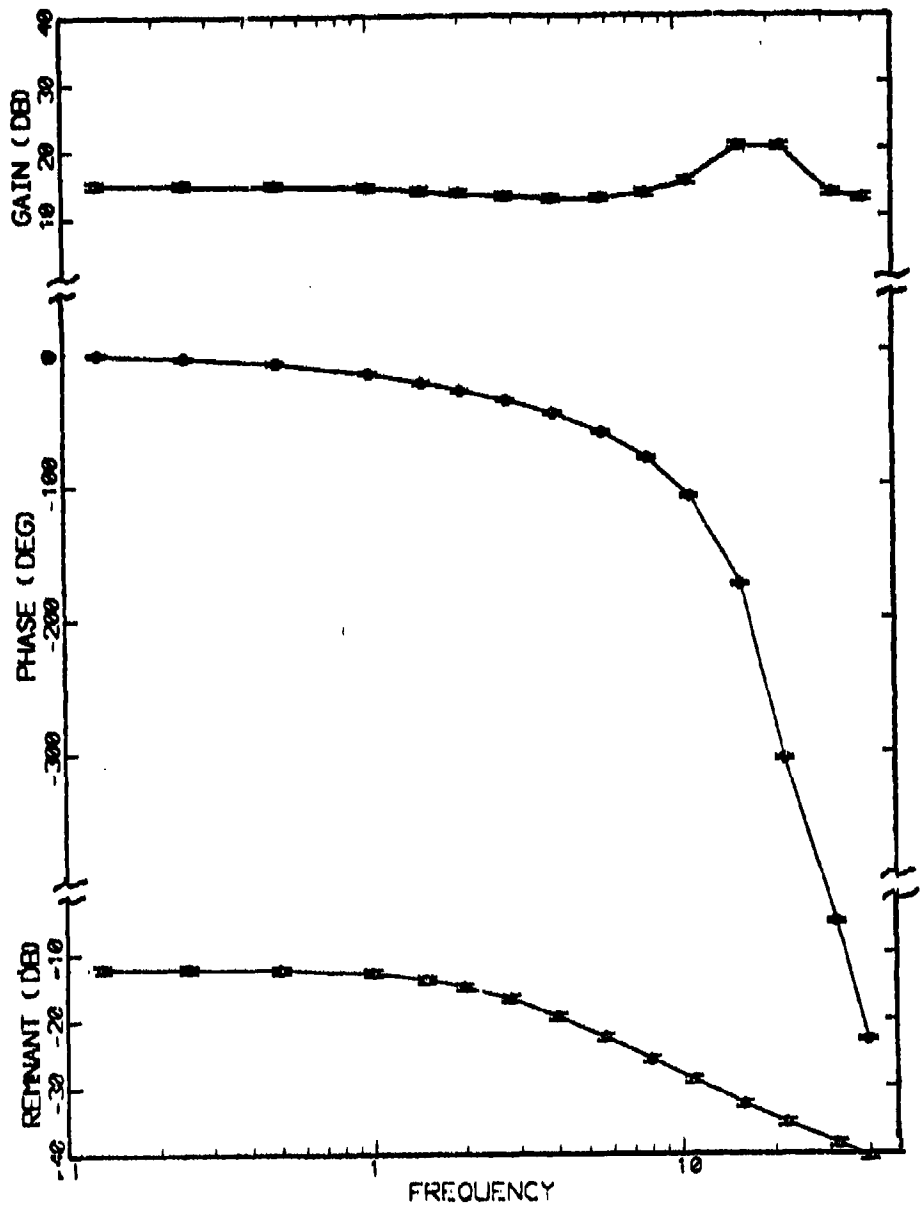


Fig. 1: Performance on Model Generated Data

Table 2: Identified Parameter and Confidence Intervals

Condition	Identified Parameters and Confidence Intervals						J
	ρ_u(dB)	ρ_e(dB)	ρ_e(dB)	ρ_u	T_D	T_N/Qr	
64 Powell							
I.G.#1,2,3	-35.5	-17.9	-21.9	-25.9	.2	.1/4.0 E-4	.084
65 Powell							
I.G.#1,2,3	-35.5	-17.9	-21.9	-25.9	.2	.1/4.0 E-4	.084
Confidence Intervals							
I.G. #1	4.68	1.02	0.56	5.7	0.0013	4.26* E-5	---
I.G. #2	4.73	1.53	0.57	5.4	0.0016	4.38* E-5	---
I.G. #3	4.17	1.1	0.50	5.4	0.0015	3.52* E-5	---
Quasi-Newton							
I.G.#1,2,3	-35.5	-17.9	-21.9	-25.9	.2	.1/4.0 E-4	.084
Confidence Intervals							
I.G.#1,2,3	5.87	1.85	0.77	6.67	.002	4.2* E-5	---

* Confidence interval for the control-rate weighting.

updates the entire hessian matrix at each iteration (and does so by taking fixed pertabations about the minimum), the confidence limits are only a function of the minimum point itself. Confidence limits using Powell's 1965 Method, on the other hand, depend on the directions used and the line search step size. Therefore, they are different for each initial guess.

The 95% confidence intervals appear to be reasonable for both methods, since the identified parameters always lie within them. This is not a very strong statement because the interpretation of the confidence limits hinges on how good our assumptions were in forming the innovation covariance (weighting matrix). Sensitivity studies also show similar results.

Table 3 shows the number of function calls required by each method. In all cases the Quasi-Newton method converges using the least number of function calls, followed by Powell's 65 and 64 methods respectively. Powell's 64 Method uses considerably more function calls than the other two methods. This would seem to indicate that the other methods are better suited to this particular problem. However, closer examination reveals that only a few iterations (2 or 3) are needed

to get very close to the minimum with the rest of the iterations used for final convergence. Therefore, this method may be good to use (with a larger stopping criterion) when far from the minimum.

Table 3: Efficiency of Methods

Initial Guesses No.	Function Calls		
	'64' Powell	'65' Powell	Q-N
1	378	189	100
2	303	71	56
3	425	235	106

If we were to monitor each of the algorithm's progress to the minimum (Figure 2 shows a typical convergence profile) we would find that:

- o The Quasi Newton method
  - usually has the fewest number of iterations,
  - There is steady reduction in J, with large decreases occurring in the first few iterations,
  - converges faster than the other methods, when close to a well defined minimum

- o Powell's 64 Method
  - usually has a large initial decrease in J
  - much slower convergence close to the minimum

- o Powell's 65 Method
  - tends to have uneven convergence, that is, it may show a large decrease in J then very little decrease for several iterations, followed by a rapid decrease,
  - uses many more iterations, but fewer function calls per iteration than the other methods.

Looking back it is clear that the Quasi-Newton's convergence speed is due to the updating of the entire hessian at each iteration. Likewise, the uneven convergence of Powell's 65 Method is due to its updating of only part of the hessian at each iteration. (i.e. when large steps are taken old information incorporated into the hessian is grossly innacurate, and several iterations are required to completely update it.)

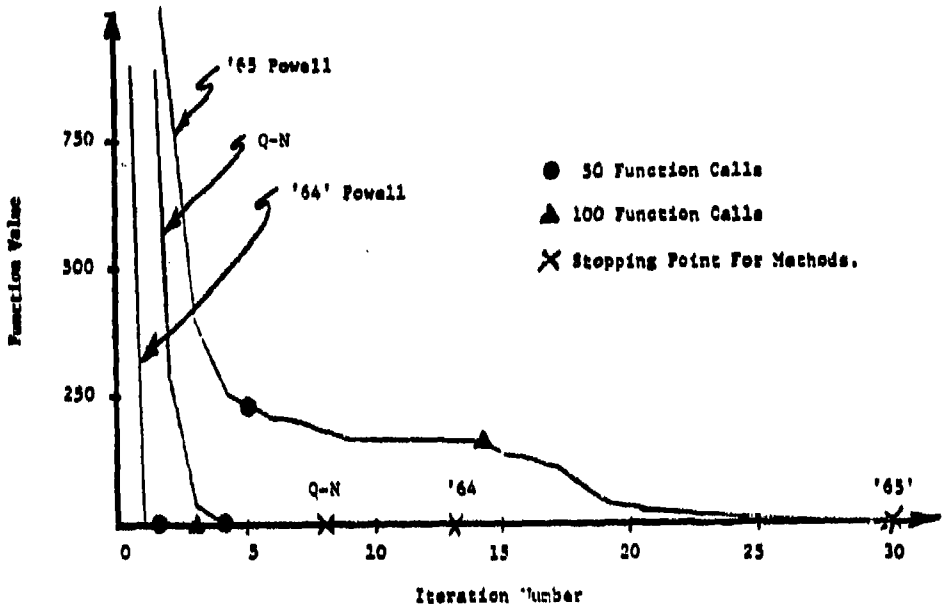


Fig.2: Convergence Profile for Initial Guess No. 1

Performance on Experimental Data

We have observed that all three algorithms work reasonably well on model generated data, with contrived sample variences. Now an example is given where the methods are applid to experimental K/s tracking data from [2]. This particular example represents a difficult problem to test the methods on. The identification is difficult because a long gently sloping valley exists in the cost surface. Starting from a good initial guess one of the parameters (noise variance on control observations) must be moved essentially to infinity, since these observations are not needed for K/s tracking [1,2,3]. All other parameters remain close to their initial values.

Table 4 shows the result of this identification, and Figure 3 presents the model's match to the data. The confidence intervals correlate well with what we would expect from sensitivity studies. All methods appear to identify the same minimum, but notice that the final function value is relatively large, because the model cannot reproduce the experimental data exactly.

Table 4: Performance on Experimental Data

Condition	Parameters and Confidence Bounds						J	Function Calls
	ρ_u	ρ_e	ρ_e	ρ_u	T_D	T_N/Q_R		
Initial guess	-40	-20	-20	-25	.17	.08	2.05	
64 Powell	-43.4	-20.5	-19.0	56.9	.16	.073/1.1 E-4	.70	103
65 Powell	-42.7	-20.7	-18.9	16.8	.16	.074/1.2 E-4	.72	256
± SD	13.5	1.45	1.98	11.5	.014	7.3 E-5		
Quasi-Newton	-43.3	-20.5	-18.9	3.86	.16	.073/1.1 E-4	.71	321
± SD	6.4	4.4	2.8	16.93	.036	1.6 E-4		

Notice in this case that Powell's 64 Method is by far the most efficient! This is a complete reversal of our findings of the previous section. This example clearly shows the advantage of using conjugate directions if a parameter must transverse a long valley in the cost surface.

These algorithms have also been applied to numerous other experimental conditions besides K/s dynamics. The following characteristics have been observed in using these methods:

- Quasi-Newton Method may develop numerical problems far from the minimum. Several initial parameter guesses may be required for convergence.

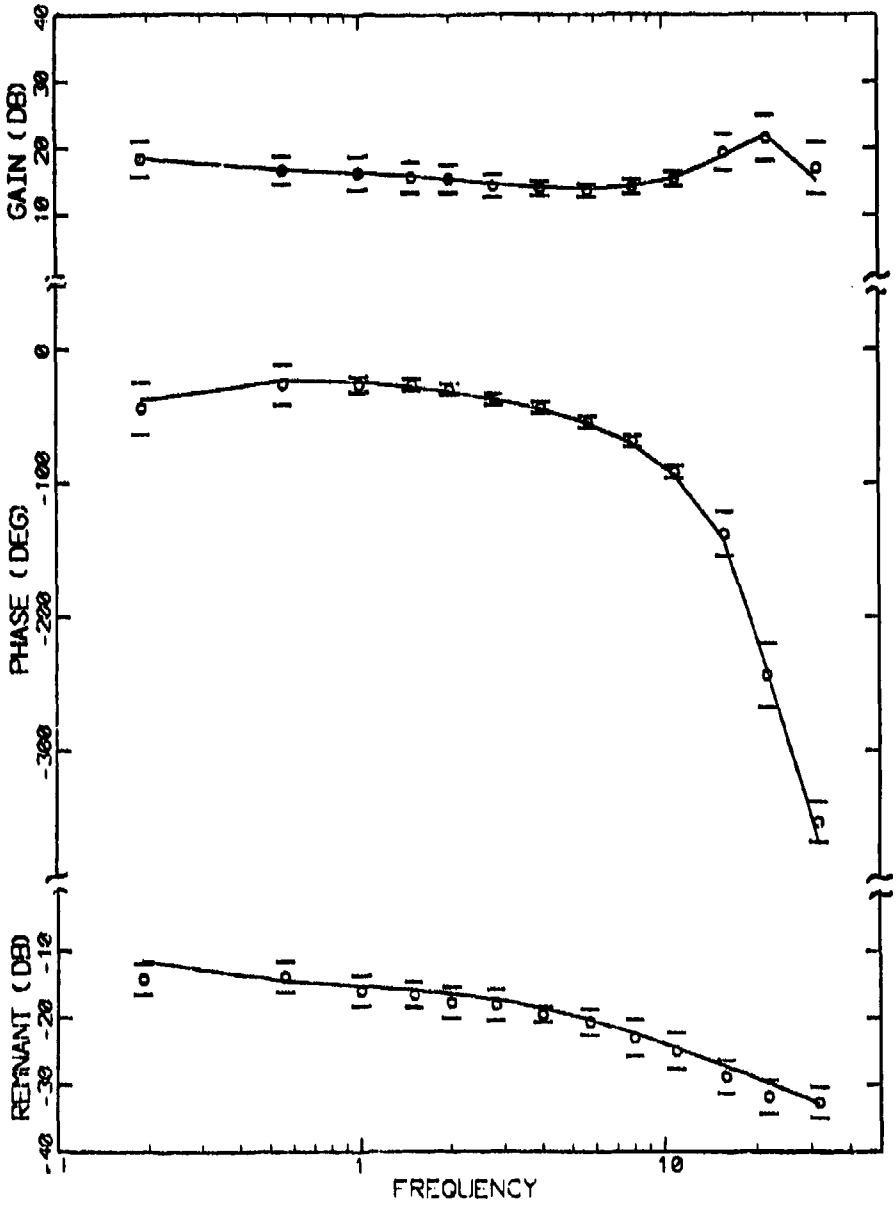


Fig. 3: Performance on Experimental Data

- All the methods may stall. This is usually due to either a stopping criterion which is too loose, or the hessian becoming singular. If the hessian does become singular, a restart usually enables it to converge.

- Of all the methods, Powell's 65 Method was found to be the most robust. In many cases it would converge without stalling, whereas the others would. This may be due to scaling the hessian at each iteration, which keeps it well conditioned

- In general, it is hard to tell which method will minimize J in the fewest number of function calls. The best approach may be to use a combination of the methods. For example;
  if close to the minimum use either:

  - Quasi-Newton alone,

  - or 64 Powell with a generous stopping criterion followed by Quasi-Newton

  if far from the minimum use either:

  - 65 Powell followed by Quasi-Newton

  - or 64 Powell followed by Quasi-Newton

Conclusions

The suitability of using function minimization techniques to identify key parameters in the steady state optimal control model was verified using model simulated and experimental data. A sum of squares objective cost function, which utilizes typical reduced order experimental data, was discussed. The choice of weightings in this cost function appeared to be an adequate description of the effective innovations covariances, and seemed to yield realistic confidence intervals, but certainly more work should be done in this area.

In general, it was found that the best minimization method to use will depend on the nature of the cost contours. However, experience using the methods seems to show that a combination of the methods will provide the best results over a wide range of cases.

This investigation has uncovered several possible areas for future research. These topics include:

- Investigating the effects (on the identified parameters) of assumptions made in forming the cost function, and accounting for correlations in the data.

- Determining sources of identifiability problems, i.e. far from the minimum, or using band-limited data.

- o Determination of better, more meaningful confidence intervals. Perhaps using a bounded-but-unknown approach and arriving at bounds on the parameters.

- o Effects of line search accuracy on the identification.

- o Advantage of adjusting the step length to minimize numerical problems when computing the forward differences.

References

1. Lancraft, R. E. and Kleinman, D. C., "A Comparison of Motor Submodels in the Optimal Control Model," 14th Annual Conference on Manual Control, April 1978.

2. Kleinman, D. L. and, S. Baron, "Manned Vehicle Systems Analysis by Means of Modern Control Theory," NASA CR-1753, June 1971.

3. Levison, W. H., Baron S., and Junker. A. M., "Modeling the Effects of Environmental Factors on Human Control and Information Processing," AMRL-TR-76-74, Aug. 1976.

4. Levison, W. H., and Junker, A. M., "A Model for the Pilots Use of Motion Cues in Roll-Axis Tracking Tasks," AMRL-TR-77-40, June 1977.

5. Schweppe, F. C., **Uncertain Dynamic Systems**, Prentice-Hall, Inc., 1973.

6. Powell, M. J. D , "An Efficient Method for Finding the Minimum of a Function of Several Variables Without Calculating Derivatives," **Computer J.**, Vol. 7, 1964.

7. Powell, M. J. D., "A Method for Minimizing a Sum of Squares of Non-linear Functions without Calculating Derivatives," **Computer J.**, 1965.

8. Himmelblau, D. M., **Applied Non-linear Programming**, McGraw-Hill, 1972.

Acknowledgements

This work is a continuation of that originally begun at The University of Conn., and presented as a short paper at last year's Annual conference on Manual Control, by Gordon Lee and Bjorn Wittenmark.

This work was supported in part under AFOSR contract No. F49620-78-c-0073.

IDENTIFICATION OF MULTIVARIABLE DYNAMIC BEHAVIOR OF MAN BY A CORRELATION TECHNIQUE*

Bernhard Tilemann

Institut fUr Regelungstechnik (Prof.Dr.-Ing. H. Rake)
RWTH Aachen, Templergraben 55, D-5100 Aachen, FRG

ABSTRACT

Investigations of the dynamic behavior of the eye-hand transmission channel of man in multivariable systems are performed by means of a correlation technique and Fourier analysis. A stochastic disturbance input is used for the identification of the human controller in the closed loop. First results of experiments with processes containing second-order controlled elements such as first-order lags and integrators with first-order lags are presented. The influence of cross-coupling effects in the human controller is discussed.

1. INTRODUCTION

Dynamic behavior of man has been examined under different aspects in the last 15 years. In most of the investigations forcing functions were used which only allowed the determination of the operator describing functions at well defined frequencies (Refs. 1-5).
An overall view of the frequency characteristics can be obtained, if forcing functions with continuous spectral densities are used. Due to human information processing limitations and for technical reasons, the spectra must have limited bandwidths. These conditions are satisfied by low-pass filtered white noise. Investigations with this type of forcing function can for example be performed by means of correlation techniques or parame-

* The investigations were performed on a hybrid computer of the "Deutsche Forschungsgemeinschaft" (DFG).

ter estimation procedures. The latter method presumes that there exists a model (of not too high order) whose parameters can be estimated (Ref. 6). On the other hand, application of a correlation and Fourier transformation procedure permits the calculation of discrete values of describing functions without postulating any model in advance. This fact seems to be convenient for the investigation of cross-coupling effects in the human controller - in particular, if the controlled elements do not contain any cross-coupling signal paths.

2. MATHEMATICAL MODEL

Although the human controller dynamics are non-linear and time-variant in general, they can be approximated by a quasi-linear model H(p) and a remnant noise n(t), if the controller is well-trained and if he performs simple control-tasks (Refs. 1,2,6). For a given two-dimensional control task (Fig.1)

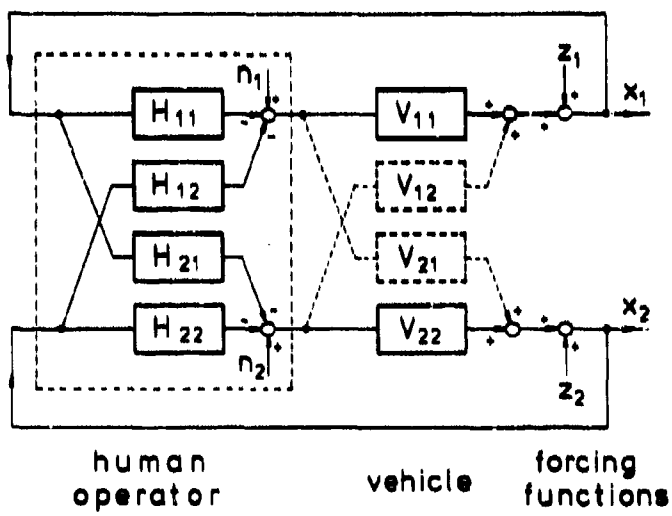


Fig. 1 Two-variable control loop
(with uncoupled controlled elements)

this model shall be represented in terms of the 2x2-matrix $\underline{H}(p)$ of the controller describing functions.

We consider the closed multivariable control loop of n^{th} order (Fig.2),

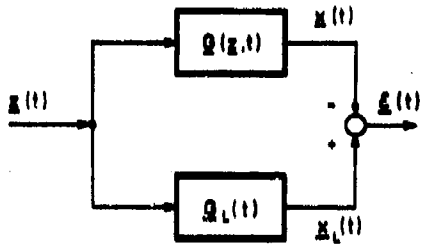


Fig. 2 Model error $\varepsilon(t)$

whose nonlinear dynamic properties are described by the matrix $\underline{G}(\underline{z}(t),t)$. A quasi-linear multivariable system with the matrix $\underline{G}_L(t)$ of its impulse responses shall be determined, so that the cost functional

$$q(\varepsilon) := \sum_{i=1}^{n} \overline{\varepsilon_i^2(t)} = \sum_{i=1}^{n} \overline{(x_{Li}(t) - x_i(t))^2} \qquad (1)$$

becomes minimal. These conditions lead to a system of Wiener-Hopf integral equations

$$\underline{\Phi}_{zx}(\tau) - \int_0^{\infty} \underline{G}_{Lo}(t) \cdot \underline{\Phi}_{zz}(\tau - t)\, dt = 0 \qquad (2)$$

as implicit calculation rule for the matrix $\underline{G}_{Lo}(t)$ of the optimal impulse responses (Ref.7). In Eq.(2) $\underline{\Phi}_{zx}$ and $\underline{\Phi}_{zz}$ are the matrices of the cross- and auto-correlation functions respectively. The solution of Eq.(2) in the frequency domain is

$$\underline{F}(p) = \mathcal{F}\{\underline{G}_{Lo}(t)\} = \underline{S}^t_{zx}(p) \cdot [\underline{S}^t_{zz}(p)]^{-1} . \qquad (3)$$

$\underline{F}(p)$ with $p = j\omega$ is the matrix of the frequency responses of the whole multivariable control system (e.g. represented by Fig.1).

It is determined by $\underline{F}(p) = [\underline{I} + \underline{F}_o(p)]^{-1}$ (4a)

and $\underline{F}_o(p) = \underline{V}(p) \cdot \underline{H}(p)$ (4b)

with $\underline{F}_o(p)$ and $\underline{V}(p)$ being the matrices of the open-loop and the controlled element frequency responses respectively. From Eqs. (3) and (4) the unknown matrix $\underline{H}(p)$ of the human controller describing functions can be calculated in general form:

$$\underline{H}(p) = \underline{V}^{-1}(p) \cdot \left\{ \underline{S}^t_{zz}(p) \cdot [\underline{S}^t_{zx}(p)]^{-1} - \underline{I} \right\} . \quad (5)$$

For the control loop configuration shown in Fig.1 the first element of $\underline{H}(p)$ is given by

$$H_{11} = \frac{1}{\det(\underline{V}) \cdot \det(\underline{F})} [V_{22} \cdot (F_{22} - \det(\underline{F})) + V_{12} \cdot F_{21}]. \quad (6)$$

Similar expressions can be obtained for the remaining three describing functions H_{12}, H_{21}, and H_{22}. For notational convenience, the argument p has been omitted in all terms of Eqs. (6) to (10). The calculation of the elements of the matrix $\underline{F}(p)$ is demonstrated by one example:

$$F_{22} = \frac{1}{\det(\underline{S}_{zz})} [S_{z2x2} \cdot S_{z1z1} - S_{z1x2} \cdot S_{z2z1}] . \quad (7)$$

The cross-power spectral densities $\underline{S}_{zx}$ and $\underline{S}_{zz}$ are obtained by Fourier transformation from the corresponding cross-correlation functions $\underline{\phi}_{zx}$ and $\underline{\phi}_{zz}$ (Ref.8), e.g.

$$\phi_{z1x2j} = \frac{1}{N-j+1} \sum_{i=0}^{N-j} z1_i \cdot x2_{i+j} , \quad j=0,1,\ldots,M, \quad (8)$$

and

$$\mathrm{Re}(S_{z1x2k}) = T \cdot \left[\phi_{z1x2o} + 2 \sum_{j=1}^{M-1} \frac{1}{2} (\phi_{z1x2j} + \phi_{x2z1j}) \cos \frac{kj}{M} \pi + \right.$$

$$\left. + \frac{1}{2} (\phi_{z1x2M} + \phi_{x2z1M}) \cos k\pi \right] , \quad (9)$$

$$Im(S_{z1x2k}) = T\cdot\left[2\sum_{j=1}^{M-1}\frac{1}{2}(\phi_{z1x2j} - \phi_{x2z1j})\sin\frac{kj}{M}\pi\right], \qquad (10)$$

$$k=0,1,\ldots,M,$$

where $z1_i = z1(i\cdot T)$ and $x2_{i+j} = x2((i+j)T)$ represent the sampled data, N+1 is the number of data-points, and M+1 is the number of points of the correlation and spectral density functions.

3. EXPERIMENTAL DESIGN

Two noise generators produce pseudo-random gaussian noise signals with power spectral densities constant for frequencies $f < 50$ cps (314 rad/s). These signals are fed into two second-order low-pass filters with 0.2 cps (1.26 rad/s) bandwidth which provide the forcing functions z_1 and z_2. The controlled elements are simulated on an analog computer which is part of a hybrid computer system.*) The human controller controls the system by means of a two-dimensional joy-stick without spring forces, almost free of mass and friction. The controlled variables x_1 and x_2 are shown to the operator by means of an oscilloscopic display. They are represented as the cartesian co-ordinates of a small circle with the system of co-ordinates indicated by a vertical and a horizontal line (Fig.3).

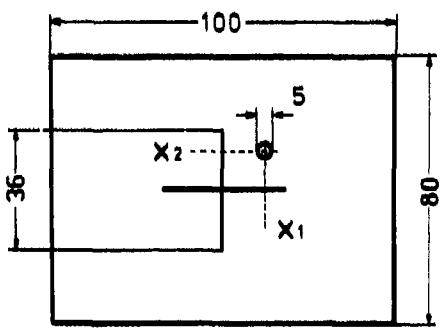


Fig. 3 Signal representation
(measures in millimeters)

*) Analog computer: DORNIER DO 720, digital computer: RXDS Sigma 3

The subjects under test sit in a distance of 80 cm from the display and actuate the stick in order to minimize the controlled variables x_1 and x_2, i.e. they try to transfer the circle into the origin of the co-ordinate system. One test of this kind is usually run for about 3 minutes.

The describing functions of the human controller are derived from recordings of the forcing functions and the controlled variables. Each of these four signals is partitioned into two frequency intervalls

1. $0.042 \text{ cps} \leq f^L < 1\text{cps}$ ($0.26 \text{ rad/s} \leq \omega < 6.28 \text{ rad/s}$) and
2. $0.333 \text{ cps} \leq f^H < 10 \text{ cps}$ ($2.09 \text{ rad/s} \leq \omega < 63.8 \text{ rad/s}$)

by appropriate analog filters (Ref.9). The filtered signals are sampled at rates of $T^L = 120$ ms and $T^H = 15$ ms in the lower and higher frequency band respectively. This corresponds to Shannon frequencies of $f^L_{Sh} = 4.16$ cps (26.1 rad/s) and $f^H_{Sh} = 33.3$ cps (209.4 rad/s). As each correlation function has a length of $M+1 = 100$ data points, the describing function values to be calculated are equally spaced with

$\Delta f^L = 0.042$ cps in the lower frequency range and
$\Delta f^H = 0.333$ cps in the higher frequency range.

In general the functions are plotted from 0.042 to about 5 cps (31 rad/s) and contain 23 points in the lower and 15 points in the higher intervall.

Partitioning of signals as described gives the possibility to amplify the higher frequency signals by a factor 40 prior to sampling and A/D-conversion and thus to improve their evaluation. As the signals z_1, z_2, x_1, and x_2 are treated in the same manner, and as only quotients of power spectral densities are used (Eq.7), the final results are not affected by this manipulation.

Fig.4 shows the influence of band-pass filtering on the forcing function characteristics.

The analog signals are digitized via analog-digital-converters and are stored on magnetic tape. The execution of the control tasks and the data-sampling is supervised by the digital component of the hybrid system, while the signal processing is done off-line with a larger computer of the computing center of the RWTH Aachen*).

*) Control Data Cyber 175

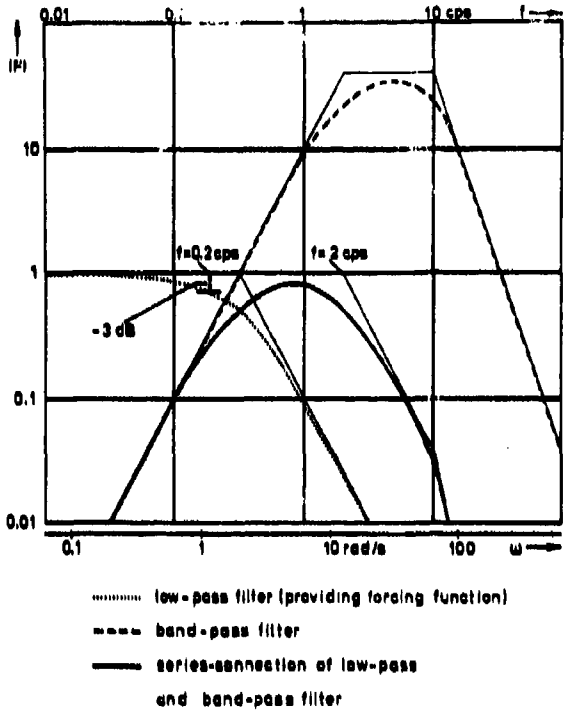


Fig. 4 Influence of band-pass filtering on forcing function characteristics

4. EXPERIMENTAL RESULTS

During the experiments, vehicle dynamics consisting of first-order lags and integrators with first-order lags were simulated on the analog computer. In a first series of experiments, the time-constants of the first-order lags were varied, in a second one the integrating rates of the integrators. The dynamics in each control loop were identical, i.e.

$$V_{11}(p) = V_{22}(p) \text{ with } V_{12}(p) = V_{21}(p) = 0.$$

The functions shown in Figs. 5 and 6 are averages of results from two tests conducted with different subjects.

Results with first-order lags

The frequency function of the lag term is

$$V_{11}(p) = V_{22}(p) = \frac{K_V}{1+pT} . \qquad (11)$$

The gain K_v is fixed to 5.85 cm/Degr. for all control tasks, while the time constant T is set to 0.3, 0.5, and 1.0 s.
The open-loop describing functions F_{o11} and F_{o22} have an approximately integral character corresponding to the "Crossover Model" of McRuer et al. (Ref.1). The related average crossover frequency is f_c = 1.8 cps (11.3 rad/s), and the phase margin is $\varphi_M = 65^o$.
The corresponding controller describing functions H_{11} and H_{22} can be interpreted as proportional-integral terms (Fig. 5).
The magnitudes of H_{11} and H_{22} depend on the time constant T varied during the tests. This is to be seen most clearly in the vicinity of the crossover frequency f_c. For frequencies f > 3 cps (19 rad/s), the magnitudes as well as the phases show a high variability, so that interpretation of the curves is impossible. The phases of H_{11} and H_{22} do not seem to depend on T. They differ little over a wide frequency range and show the influence of a time delay.
All those observations can be expressed by the following models:

$$H_{ii}(p) = K_H\left(1 + \frac{1}{pT_I}\right)\cdot e^{-pT_D} \quad , \quad i=1,2, \qquad (12)$$

where K_H is the model gain, T_I is the model integrating time constant, and T_D is the delay time constant. This model describes the experimental data fairly well in the range of 0.1 cps<f<3 cps (0.63 rad/s < ω < 19 rad/s), but it does not account for the very low frequency data (f < 0.1 cps). Prior to the extension of the model to this frequency region more data must be accumulated. Then also the values of the model parameters can be determined satisfactorily.

The describing functions of the human controller cross-coupling terms H_{12} and H_{21} show no significant dependency on the vehicle dynamics. They vary strongly and thus permit the assumption that at least for frequencies f > 0.3 cps (1.9 rad/s) no cross-coupling between the two control loops exists. In this region the mean values of $|H_{12}|$ and $|H_{21}|$ are about 20 dB smaller than

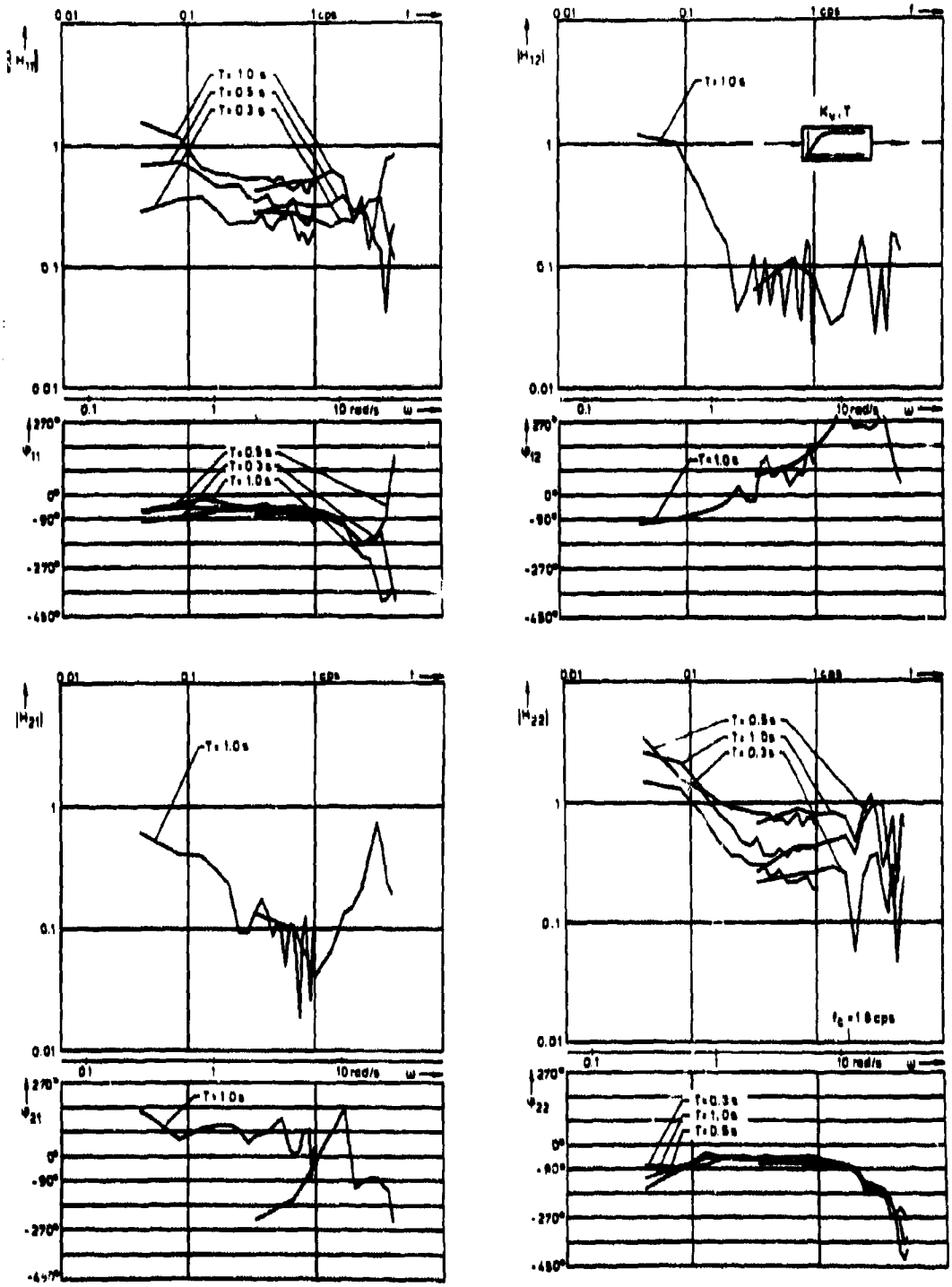


Fig. 5 Human controller describing functions
(with first-order lags)

those of $|H_{11}|$ and $|H_{22}|$. For frequencies f < 0.3 cps there might exist a slight coupling between the loops. This could be attributed to low-frequency corrections of those operator control commands, which do not exactly go into the desired "direction", as can be observed on the oscilloscopic display. For reasons of clarity, the describing functions of H_{12} and H_{21} have been plotted only for T = 1.0 s in Fig. 5.

Results with integrators and first-order lags

In a second series of experiments the controlled element dynamics were of the form

$$V_{11}(p) = V_{22}(p) = \frac{K_I}{p} \cdot \frac{K_V}{1+pT} \quad . \tag{13}$$

The time constant T=0.1 s and the gain K_V=7.8 cm/Degr. are constant for all control tasks, while the integrating rate K_I is varied with 0.5, 0.9, and 1.3 s^{-1}.
The open-loop describing functions F_{o11} and F_{o22} again show integral character with an average crossover frequency f_c=1.6cps (10.0 rad/s) and a phase margin $\varphi_M = 50^o$. The values for f_c and φ_M are somewhat smaller than those obtained with first-order lags. This means that the stability reserve is smaller in the second case, i.e. the control loops are nearer to the stability margin and thus more difficult to control. This statement is confirmed by corresponding subject-ratings. The controller describing functions H_{11} and H_{22} can be approximated by models of the form

$$H_{ii}(p) = K_H(1+pT_L)\cdot e^{-pT_D}, \quad i=1,2, \tag{14}$$

where K_H and T_D are defined as in Eq.12, and T_L is the model lead time constant. The Bode-plots in Fig.6 show that the gains K_H of H_{11} and H_{22} increase with decreasing integrating rate K_I, while the phases do not seem to depend on K_I. As with first-order lag controlled elements, the models fit well in the region of the crossover frequency, but they do not describe the expe-

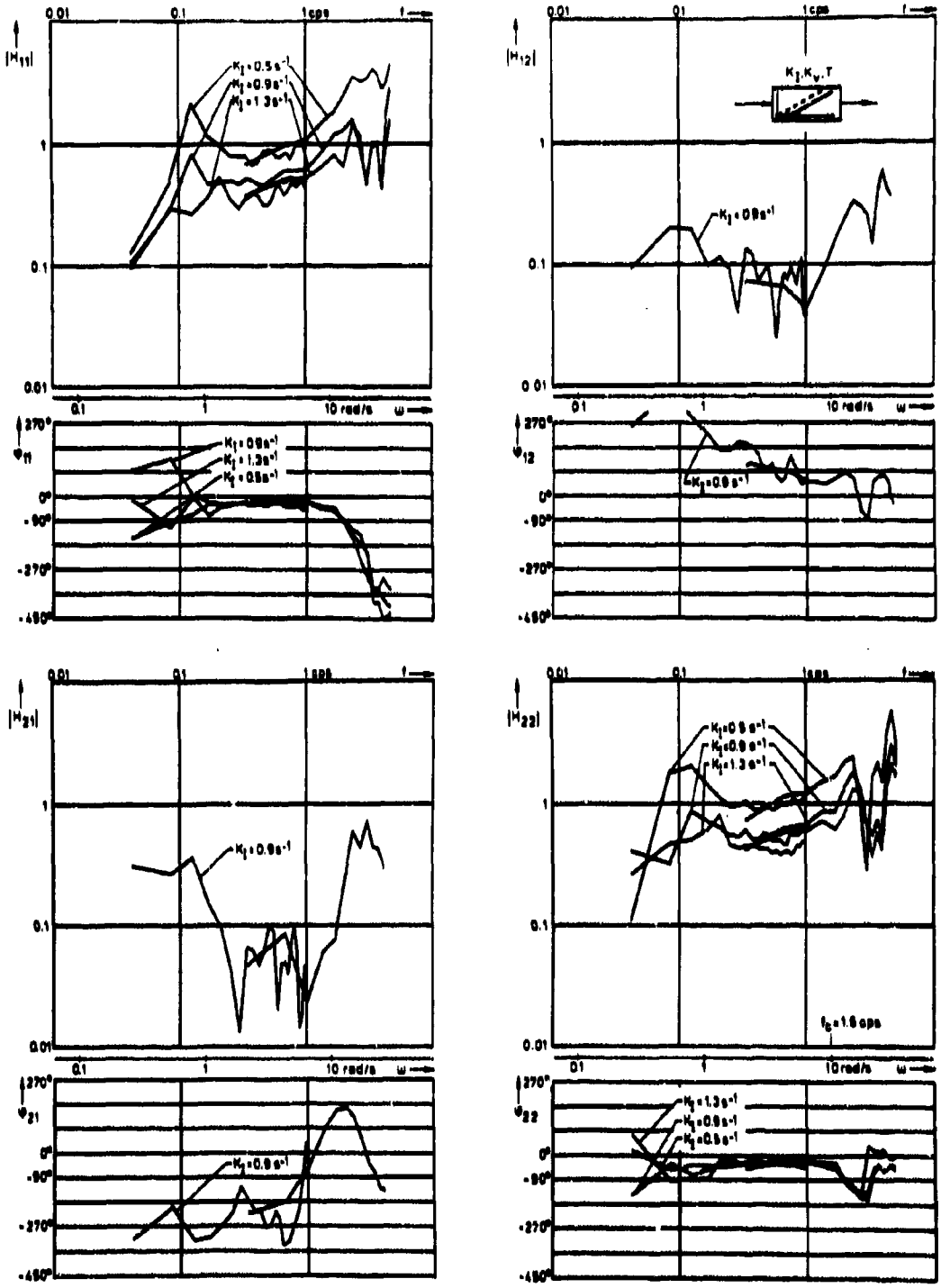


Fig. 6 Human controller describing functions (with integrators and first-order lags)

rimental data for frequencies below 0.2 cps (1.25 rad/s) and over 2.5 cps (15.7 rad/s).

A last attempt to interpret the cross-coupling influences in the human controller dynamics consists in the consideration of the range in which all magnitudes of all coupling terms H_{12} and H_{21} for both types of controlled elements are to be found (Fig.7). No dependency on the vehicle dynamics can be observed. All calculated values fall into the dotted area. There might be a correlation between the bandwidths of the forcing functions (0.2cps) and the coupling term magnitudes, because for frequencies $f > 0.2$ cps the magnitude values are significantly smaller than for lower frequencies. For further investigation of these cross-coupling influences partial coherence functions will be calculated.

5. CONCLUSIONS

Results of investigations of human controller dynamics in two-axis control tasks by means of a correlation technique have been presented. The experimental data have been interpreted by quasi-linear models in the frequency domain. The presence of cross-coupling terms in the models in the case of uncoupled controlled elements could not be proved, but the calculatod cross-coupling describing functions seem to be independent of the vehicle dynamics. Further investigation of these influences, for instance by means of coherence functions, is intended.

REFERENCES

1. McRuer, D.T.; Graham, D.; Krendel, E.S.; Reisener, W.Jr.: Human Pilot Dynamics in Compensatory Systems. USAF Flight Dynamics Lab., Wright-Patterson-AFB, Ohio,Techn. Rep. AFFDL-TR-65-15, July 1965.

2. Levison, W.H.; Elkind,J.I.: Studies of Multivariable Manual Control Systems - Two-Axis Compensatory Systems With Separated Displays and Controls. NASA CR-875, Oct. 1967.

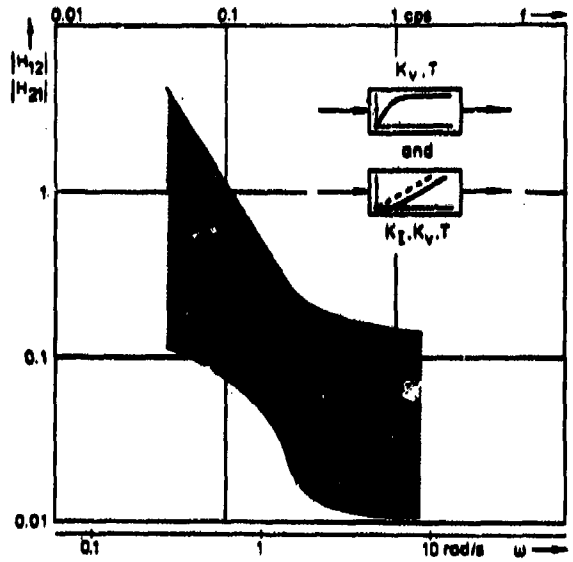


Fig. 7 Range of all magnitudes $|H_{12}|$ and $|H_{21}|$

3. Levison, W.H.; Elkind, J.I.; Ward, J.L.: Studies of Multivariable Manual Control Systems - A Model for Task Interference. NASA CR-1746, May 1971.

4. Jex, H.R.; Jewell, W.F.; Allen, R.W.: Development of the Dual-Axis and Cross-Coupled Critical Tasks. 8th Annual Conference on Manual Control, Ann Arbor, Michigan, USA, 1972.

5. Ritchie, M.L.; Nataraj, N.S.: Human Performance Evaluation in Dual-Axis Critical Task Tracking. 11th Annual Conference on Manual Control, 1975.

6. Todosiev, E.P.; Rose, R.E.; Summers, L.G.: Human Performance in Single- and Two-Axis Tracking Systems. 2nd Annual NASA-University Conference on Manual Control, MIT, NASA SP-128, 1966.

7. Schwarz, H.: Mehrfachregelungen I. Verlag Springer, Berlin, Heidelberg, New York, 1967.

8. Rake, H.: Identifizierung linearer Systeme mit statistischen Methoden. Inaugural dissertation, RWTH Aachen, FRG, 1969.

9. Kuchen, B.: Das Übertragungsverhalten des Menschen im geschlossenen Regelkreis bei stochastischen Testsignalen. Ph.D.Thesis, RWTH Aachen, FRG, 1974.

TRANSFER FUNCTION PROPERTIES OF THE OPTIMAL CONTROL MODEL IN TARGET-TRACKING TASKS

by
Anil V. Phatak, Senior Engineer
Analytical Mechanics Associates
2483 Old Middlefield Way
Suite 210
Mountain View, California 94043
and
Kenneth M. Kessler, Senior Engineer
Systems Control, Inc. (Vt)
1801 Page Mill Road
Palo Alto, California 94304

ABSTRACT

The optimal control model for the human operator has been used to describe human performance in target-tracking tasks as in air-to-air and AAA situations. Regardless of the specific formulation used, an internal model (to the operator) of the target trajectory, control system dynamics, and human psychophysical limitations is required. Typically, the target-trajectory is assumed to be represented by a polynomial in time (i.e. constant target velocity or acceleration assumption).

This paper analyzes the properties of the equivalent input/output operator transfer function corresponding to the optimal control model. In the absence of time delay and motor noise, the zeros in the operator transfer function exactly cancel the poles of the assumed control system dynamics. This result is extremely useful since it allows for a simpler representation of the optimal control model.

INTRODUCTION

A systematic approach to target-tracking tasks requires the development and integration of models of the control system dynamics, the target trajectory, and the human operator. These models can be incorporated into a composite analysis algorithm that can be used for analytical and predictive purposes. All too often, system performance is evaluated by utilization of the standard optimal control model [1] which requires explicit state vector representation of the plant/noise dynamics (control system/target trajectory). The output-error vector model formulation (the proportional-integral-derivative (P-I-D) structure) described in this paper represents an attempt to simplify the resulting analysis algorithm by recognizing the nonuniqueness of the state representation in terms of the input/output transfer function associated with the human operator. This P-I-D model structure has been successfully used to evaluate system performance in both anti-aircraft artillery (AAA) and air-to-air-combat tracking tasks [2,3].

This work was supported by the 6570th Aerospace Medical Research Laboratory, Wright-Patterson AFB, Ohio, under Contract No. F33615-77-C-0507.

In these previous applications, the following modifications to the standard optimal control model have proved useful: (1) elimination of the time delay (and, hence, the predictor) for tracking systems; (2) combining motor noise and observation noise into one equivalent "lumped" observation noise; (3) use of an internal model (to the operator) in terms of the output (tracking) error displayed variables, alone; and (4) generalization of the quadratic cost to allow for higher order control derivatives (e.g. $\ddot{u}$), depending upon the structure of the control system and the assumed target-trajectory dynamics.

One advantage in using these modifications is the relative ease in selecting the P-I-D model parameters as functions of target-trajectory characteristics for the purpose of matching input/output operator data, typically the ensemble time histories. Note that the analysis of such systems is complicated by the time-dependent model parameters associated with the characteristics of the target trajectory.

In the subsequent development, it is shown that in the absence of time delay and motor noise, the numerator of the operator input/output transfer function exactly cancels the denominator of the assumed control-system dynamics. While the time delay in the observation is not explicitly used in the formulation, it can be incorporated implicitly into the plant/disturbance internal model by augmenting the state vector to include states corresponding to a Padé approximation for the time delay. In this case, exact cancellation occurs in the operator and control-system dynamics. Near cancellation occurs when the time delay is explicitly represented in the display while retaining a low-order Padé approximation in the operator transfer function.

MATHEMATICAL DEVELOPMENT

The general configuration of the target-tracking system is shown if Figure 1. In this figure, $\theta_T(t)$ is the commanded (target) trajectory, $\theta(t)$ is the output of the control system, $e(t)$ is the tracking error, and $T(s) \triangleq \theta(s)/u/(s)$ is the control-system transfer function.

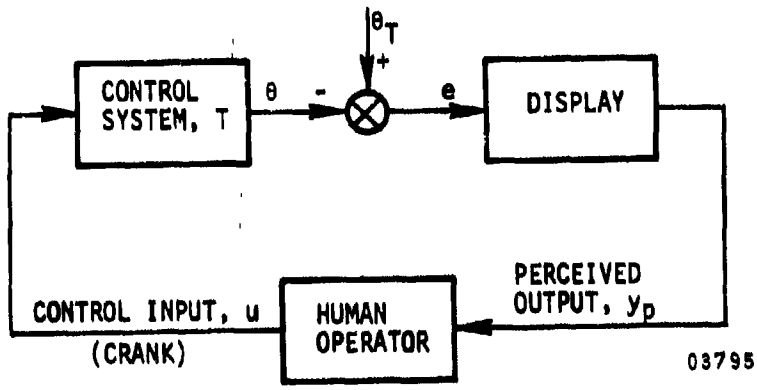


Figure 1 Manual Target-Tracking System

This system is representative of an assumed single input, single output (SISO) where it is further assumed that each axis is uncoupled, e.g. azimuth and elevation axes for the AAA tracking task. Consider the SISO plant/disturbance dynamics written in the form:

$$\dot{x}_a = F_a x_a + G_a u + \Gamma_a w \tag{1}$$

$w \sim N(0, W_a)$, scalar white noise

x_a = (nx1) state vector and u is a scalar

and

$y_a = H_a x_a$ be the scalar displayed output.

Let the assumed internal model for the plant/disturbance dynamics be given by

$$\dot{x} = Fx + Gu + \Gamma w \tag{2}$$

where, as before, u and w are scalar variables and

$w \sim N(0, W)$ and x = (mx1) vector, $m \leq n$ with the displayed output

$y = Hx$

The operator perceives

$$y_p = y + v_y$$

where $v_y \sim N(0, V_y)$ is the (lumped) observation noise vector.

The optimal control human operator model is given by a Kalman filter in tandem with an optimal controller in the form

$$\dot{\hat{x}} = F\hat{x} + Gu + K(y_p - H\hat{x}); \quad \hat{x}(0) = \hat{x}_0 \tag{3a}$$

$$u = -\lambda\hat{x} \tag{3b}$$

where the control and filter gains, λ and K, respectively, are obtained by solving the appropriate Riccati equations.

Figure 2 shows the optimal control operator model for the target-tracking task.

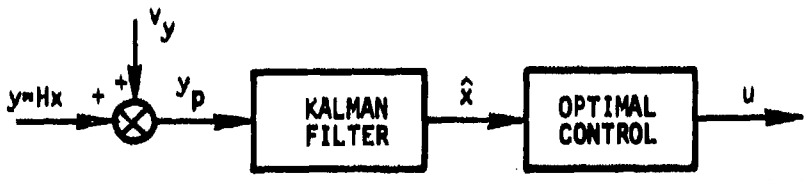


Figure 2 Optimal-Control Operator Model

The equivalent SISO operator transfer function is given by

$$\frac{u(s)}{y_p(s)} = -\lambda(sI-F+G\lambda+KH)^{-1}K = \frac{-\lambda\ \mathrm{adj}(sI-F+G\lambda+KH)K}{\det(sI-F+G\lambda+KH)} \tag{4}$$

where $\lambda \equiv (1xm)$; $K \equiv (mx1)$; $H \equiv (1xm)$; "adj" is the adjoint matrix operator; "det" is the determinant; and use has been made of Laplace transform notation for a constant coefficient system.

An alternate representation of this transfer function is to write

$$u(s) = -\lambda\hat{x}(s) = -\lambda(sI-F+KH)^{-1}[Gu(s)+Ky_p(s)] \tag{5}$$

By use of the identity

$$c(sI-A)^{-1}b = \frac{\det(sI-A+bc)-\det(sI-A)}{\det(sI-A)}$$

it is easy to show

$$\frac{u(s)}{y_p(s)} = \frac{-\lambda\ \mathrm{adj}(sI-F+KH)K}{\det(sI-F+G\lambda+KH)} \tag{6}$$

In Figure 1, let the dynamics for the commanded input be

$$\ddot{\theta}_T(t) = w(t);\quad w \sim N(0,W(t)) \tag{7}$$

corresponding to the operator's internal model of a constant-velocity target trajectory. The white noise, $w(t)$, represents his uncertainty associated with this assumption. Experimental evidence suggests that this is a reasonable assumption, however, the subsequent development is not limited to this structure; $\ddot{\theta}_T = w(t)$ can be used in those cases requiring a high degree of operator skill and low values of "cross-over" velocities. Reference 2 suggests the selection of the appropriate target model is obtained from the ensemble mean tracking error observed from data; i.e. it can be shown that the general shape of the predicted tracking error has a one-to-one relationship with $\theta_T^{(k)}$, the kth derivative of the target trajectory.

By definition, the tracking error is given by

$$e = \theta_T - \theta \text{ or } \ddot{e} = w - \ddot{\theta} \tag{8}$$

where we have used the constant-velocity target assumption.

If the internal model for the control system transfer function in Figure 1 is given by

$$T(s) = \theta(s)/u(s) = \frac{(c_{p-1}s^{p-1} + \ldots c_1 s + c_0)}{s^{\ell}(s^p + d_{p-1}s^{p-1} + \ldots d_1 s + d_0)} \tag{9}$$

where ℓ is the "type" of $T(s)$, one may write in state vector notation

$$\ddot{\theta} = H_z z \tag{10a}$$

$$\dot{z} = F_z z + G_z \tilde{u};\ z \equiv (p\times 1) \tag{10b}$$

where, for example

$$\tilde{u} = u^{(2-\ell)}(t)$$

and

$$F_z = \left[\begin{array}{c|c} 0 & I \\ \hline -d_0\ -d_1 \ \ldots\ -d_{p-1} & \end{array}\right];\quad G_z^T = [0 \ldots 0\ 1];$$

$$H_z = [c_0 c_1 \ldots c_{p-1}] \tag{10c}$$

The internal model for the plant/disturbance dynamics can be derived by augmenting Eqs. (8) and (10) as follows:

$$\text{let } e^+ = \begin{bmatrix} e \\ \dot{e} \end{bmatrix};\quad x = \begin{bmatrix} e^+ \\ z \end{bmatrix} \tag{11}$$

then

$$\dot{x} = Fx + G\tilde{u} + \Gamma w \tag{12a}$$

where

$$F = \left[\begin{array}{c|c} F_e & F_{ez} \\ \hline 0 & F_z \end{array}\right];\quad G = \begin{bmatrix} 0 \\ \hline G_z \end{bmatrix};\quad \Gamma = \begin{bmatrix} \Gamma_1 \\ \hline 0 \end{bmatrix} \tag{12b}$$

and

$$F_e = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix};\quad F_{ez} = \begin{bmatrix} 0 \\ \hline H_z \end{bmatrix};\quad \Gamma_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \tag{12c}$$

and F_z, G_z and H_z are given in Eq. (10c).

The operator is assumed to perceive only the tracking error (no rate information) of the form

$$y_p = Hx + v_y = e + v_y \tag{12d}$$

where $H = [H_e \quad 0]$; $H_e = (1 \; 0)$

Of course, the tracking error rate is estimated within the Kalman filter. The assumption of the perception of only the tracking error is a departure from the standard optimal control model and has been shown to adequately predict operator describing function and remnant in standard laboratory tracking tasks [5]. Incorporation of rate perception may be necessary to replicate observed data having sharp discontinuities in the tracking error due to high "cross-over" velocities in the target trajectory.

Task Cost Functional

The optimal control operator model requires the use of a cost functional to determine the solutions of the control and filter gains appearing in Eq. (4) or, alternatively, in Eq. (6). The cost functional is typically assumed to be of the form

$$J(u) = E\{\frac{1}{T_f}\int_0^{T_f}(x'Qx + \underline{u}'R\underline{u})dt\} \tag{13}$$

where x is given by Eq. (11), $\underline{u}' = [u, \dot{u}, \ddot{u}, \ldots u^{(k)}]$, and Q and R characterize the task requirements and control constraints, respectively. In this task, the operator attempts to achieve acceptable tracking performance with a minimum level of control effort over the duration of the tracking task, T_f.

Since one characterization of tracking performance is rms error, a reasonable Q is diagonal and of the form

$$Q \triangleq \text{diag}\,(1, 0, \ldots 0).$$

For specific application of the P-I-D operator structure, the simplest form of $\underline{u}'R\underline{u}$ is (see (Eq. (12)) $g\ddot{u}^2$, where g is the (scalar) control weight, representative of task difficulty. The selection of $\ddot{u}$ in the cost arises from the target trajectory/control system dynamics, rather than the a priori selected $\dot{u}$ as in the standard optimal control model.

In summary, the resultant cost function is of the form

$$J(u) = E\{\frac{1}{T_f}\int_0^{T_f}(e^2 + g\ddot{u}^2)dt\} \tag{14}$$

where $\ddot{u}$ is given in Eq. (10), gives rise to appropriate Riccati equations, solutions of which are λ and K.

Operator Input/Output Transfer Function

The final derivation of the operator SISO transfer function is a consequence of matrix manipulation arising from the system described in Eq. (12). Thus, if the Kalman filter gains are represented by $K' = [K_e' \mid K_z']$, where e represents the error states and z represents the augmented control-system states, it can be shown

$$K_z' \equiv (0 \ldots 0)$$

and, further, Eq. (6) reduces to

$$\frac{\Omega(s)}{y_p(s)} = \frac{-\Delta_z \lambda_e \, \mathrm{adj}(sI - F_e + K_e H_e) K_e}{\Delta_{cz}\Delta_{fe} - N(s)\{\lambda_e \, \mathrm{adj}(sI - F_e + K_e H_e)\begin{bmatrix}0\\1\end{bmatrix}\}} \tag{15}$$

where

$$\lambda = [\lambda_e \mid \lambda_z]; \tag{15a}$$

$$\Delta_{fe} \triangleq \det(sI - F_e + K_e H_e); \tag{15b}$$

$$\Delta_{cz} \triangleq \det(sI - F_z + G_z \lambda_z); \tag{15c}$$

$$\Delta_z \triangleq \det(sI - F_z) = s^p + d_{p-1}s^{p-1} + \ldots d_1 s + d_0 \tag{15d}$$

and $N(s)$ is the numerator polynomial (zeroes) of the control system dynamics associated with $H_z(sI - F_z)^{-1}G_z$. The details of this derivation are given in the appendix.

Note that the human operator transfer function has zeroes (Δ_z) that exactly cancel the poles in the control system plant dynamics. The denominator of the transfer function does not, in general, have poles at $N(s)$ (zeroes of the control system transfer function). However, for most situations there exist poles near $N(s)$ which result in near cancellation of the control system zeroes.

In the case when the control system dynamics are "fast" enough (relative to the operator's limited bandwidth), the transfer function given in Eq. (10) can be approximated by $\theta(s)/\Omega(s) \approx 1$. For this case, the operator transfer function reduces to a particularly simple structure. It can be shown (see, for example, Ref. 2) that the operator input/output transfer function reduces to a Butterworth configuration.

What makes these simple P-I-D structures particularly attractive is that the pole-zero locations of the operator transfer function can be written as explicit functions of the model "tuning" parameters, W, V_y, and g. These model parameters must be selected properly in order to replicate or predict observed human response data.

SUMMARY

The P-I-D optimal control structure has found usefulness in describing human operator performance in target-tracking tasks. The structural properties of the operator transfer function are particularly attractive since it focuses on the target-trajectory model and the model parameters rather than on the (sometimes) high-order control system dynamics. It has been shown that in the absence of time delay and motor noise, the zero of the operator transfer function exactly cancels the poles of the assumed control-system dynamics.

APPENDIX

This appendix shows the derivation of Eq. (15). From Eq. (6)

$$\frac{\tilde{u}(s)}{y_p(s)} = \frac{-\lambda\ \mathrm{adj}(sI-F+KH)K}{\det(sI-F+G\lambda+KH)}$$

The numerator can be simplified as follows:

$$-\lambda\ \mathrm{adj}(sI-F+KH)K = -(\lambda_e\lambda_z)\ \mathrm{adj}\left[\begin{array}{c|c} sI-F_e+K_eH_e & -F_{ez} \\ \hline 0 & sI-F_z \end{array}\right]\begin{bmatrix} K_e \\ 0 \end{bmatrix}$$

where $K' = [K_e' \mid 0]$ and Eq. (12b) have been used. Then,

$$\mathrm{adj}\left[\begin{array}{c|c} sI-F_e+K_eH_e & -F_{ez} \\ \hline 0 & sI-F_z \end{array}\right] =$$

$$\Delta_z\Delta_{fe}(\lambda_e\lambda_z)\left[\begin{array}{c|c} (sI-F_e+K_eH_e)^{-1} & (sI-F_e+K_eH_e)^{-1}F_{ez}(sI-F_z)^{-1} \\ \hline 0 & (sI-F_z)^{-1} \end{array}\right]$$

where Δ_z and Δ_{fe} are defined in Eqs. (15b) and (15d). The numerator can then be further simplified to:

$$\mathrm{Num} = -\Delta_z\lambda_e\ \mathrm{adj}(sI-F_e+K_eH_e)K_e$$

The denominator polynomial is given by

$$\det(sI-F+KH+G\lambda) = \det\left[\begin{array}{c|c} sI-F_e+K_eH_e & -F_{ez} \\ \hline G_z\lambda_e & sI-F_z+G_z\lambda_z \end{array}\right]$$

$$= \Delta_{cz}\Delta_{fe}\ \det[I+F_{ez}(sI-F_z+G_z\lambda_z)^{-1}G_z\lambda_e(sI-F_e+K_eH_e)^{-1}]$$

$$= \Delta_{cz}\Delta_{fe}[1+\{\lambda_e(sI-F_e+K_eH_e)^{-1}F_{ez}(sI-F_z+G_z\lambda_z)^{-1}G_z\}]$$

(using the identify $\det(I+ab^T) = 1+b^Ta$. Thus, $\tilde{u}(s)/y_p(s)$ can now be written in the form:

$$\frac{-\Delta_z \lambda_e \operatorname{adj}(sI-F_e+K_eH_e)K_e}{\Delta_{cz}\Delta_{fe} + \lambda_e \operatorname{adj}(sI-F_e+K_eH_e)F_{ez} \operatorname{adj}(sI-F_z+G_z\lambda_z)G_z}$$

Theorem:

Let $H_z(sI-F_z)^{-1}G_z = \frac{N(s)}{D(s)} = \frac{N_1(s)}{\Delta_z}$

and

$$H_z(sI-F_z+G_z\lambda_z)^{-1}G_z = \frac{N_2(s)}{\Delta_{cz}}$$

Then

$$N_1(s) = N_2(s)$$

(For a proof, see Ref. 4.)

Also, note that

$$F_{ez} \operatorname{adj}(sI-F_z+G_z\lambda_z)G_z = -\begin{bmatrix} 0 \\ \hline H_z \operatorname{adj}(sI-F_z)G_z \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} N(s)$$

These last equations can be combined to produce Eq. (15).

References

1. D. Kleinman, S. Baron, W.H. Levison, "A Control Theoretic Approach to Manned-Vehicle Systems Analysis," IEEE Trans. Auto. Control, AC-16, pp.824-832, 1971.
2. A.V. Phatak and K.M. Kessler, "Modeling the Human Gunner in an Anti-Aircraft-Artillery (AAA) Tracking Task," Human Factors, 19(5), pp.477-494, 1977.
3. K.M. Kessler and A.V. Phatak, "Modeling the Human Pilot During the Tracking Phase of an Air-to-Air Engagement," Final Report under Contract No. F33615-76-C-5011, Aerospace Medical Research Laboratory (AMRL/EMT), WPAFB, Ohio, September 1976.
4. H. Kwakernaak and R. Sivan, Linear Optimal Control Systems, p.273, John Wiley, 1972.
5. Phatak, A.V., "Investigation of Alternate Human Operator Optimal Control Model Structures," 15th Annual Con. Manual Control, Dayton, Ohio, March 1979.

ESTIMATION OF PARAMETERS CHARACTERIZING HUMAN AUTONOMIC RESPONSES TO STRESS*

Paul Milgram and John Senders

Dept. of Industrial Engineering
University of Toronto, Toronto, Canada

ABSTRACT

A unique psychophysiological experiment will be described, wherein volunteer human subjects are presented with a localized, continuously fluctuating, low-pass, randomly amplitude modulated, noxious electric current and various elicited autonomic responses, such as skin potential, skin conductance and heart rate, are monitored. Treating the data as continuous time series, describing function models of the lumped input-output autonomic stress response system are identified and their parameters are estimated. It is first demonstrated that it is possible to "drive" the responses in a non-decrementing fashion, thus justifying the application of quasi-linear and quasi-stationary mathematical techniques. Results of conventional windowed periodogram spectral analysis will be compared with those of autoregressive moving average models. A parsimonious set of parameters is presented, whose role is to characterize individual responses. Potential applications include personnel selection, psychiatric diagnosis, and stress and workload analysis.

1. INTRODUCTION

In this paper we shall present a preliminary investigation of the responsivity of the human autonomic nervous system. Although not directly involved in the execution of any closed loop manual control functions, this "control system" is still an integral part of that control loop, as it is descriptive of the fluctuating autonomic state of the human operator. The influences of various environmental and psychological stresses on the performance of the "man-in-the-loop" are of obvious importance to the practitioner of manual control modelling. What we shall consider of interest here is the identification of a set of descriptive parameters which may describe a lumped model of the human autonomic nervous system (ANS) responding in a state of intense continuous stress, induced, in the laboratory, by a continuously applied noxious local electric current. In other words, our objective is to present a *data model* which may characterize, idiosyncratically, the responses of a subject to our controlled stressor, as opposed to a *process model*, whose objective would be to model the structure of the underlying psychophysiological stimulus-response system. Some overlapping of these two purposes is, of course, inevitable.

*This work was supported by National Research Council of Canada Grant No. A8839.

A schematic model of the stimulus-response system, as we treat it here, is given in Fig. 1. The placement of three blocks within the one large block is to be construed as two levels of representing the model. The single block input-output system is the one that will be dealt with in this paper; however, we shall first discuss the model in light of the three subsystems. The system input, labelled "physical stimulus", is in this case a dermally administered local electric current. It could also (however, not 'equally as well', as we shall see shortly) have been an intense fluctuating visual, thermal, auditory, or other tactile stimulus. The problem we face is that the *actual input* to the ANS is not electric current, nor even *pain*, but the *perceived aversiveness* of the stimulus, taken at any moment in time. While we obviously can not monitor this input continuously, we can produce some estimate of a psychophysical function relating subjective sensation to stimulus current. This function is labelled H_1 in Fig. 1. In fact, previous studies have shown that such a power function has an exponent close to 2 [1], implying that the subject's perception of pain is almost linearly related to the level of stimulus power. Therefore pain, as produced by the application of electric current, emerges as the most 'controllable' of the candidate stimuli, since we may, to a good approximation, govern and monitor the intensity of the real input to the ANS.

The input labelled "sensory noise" is assumed to be small and is attributed to localized dermal changes at the stimulus electrode site and to adaptation of the pain receptors over time. More crucial is the input labelled "perceptual noise", which represents such intangibles as subject fatigue, motivation, arousal, shifting of attention, etc., which collectively act to alter the subject's perception of the applied stimulus and, eventually, to habituate to it. For example, subjects tend initially to be apprehensive of the prospect of being given an 'electric shock'; however, as the experiment progresses, their perception of its aversiveness often subsides.

Shown at the ANS (H_2) output is a set of parallel outputs elicited by the "effected response". Among the many autonomically mediated transducable outputs are electrodermal responses (skin potential, skin conductance, and skin resistance), heart rate, blood pressure, pulse volume, respiration, EEG, pupil dilation, etc. Each of these has its own "noises", since the modalities and mechanisms of transduction are different. By way of example, the reciprocal of skin conductance will not render the same signal as continuously recorded skin resistance, and both of these *exogenous* measures will differ from *endogenous* recordings, such as skin potential. Furthermore, all outputs will be mediated by other than stimulus specific influences, labelled "unrelated central inputs", representing metabolic and homeostatic influences not directly attributed to the applied stimulus.

Since there is no evident mechanism through which autonomically elicited responses may influence the apparent aversiveness of the applied stress, we treat this as an *open-loop* identification problem. The assumption of "open-loopedness" may be countered by various speculatory feedback mechanisms, such as alterations in the physical characteristics of the dermal nocireceptors due to elicited electrodermal changes, or more central influences on the perceived intensity of the stimulus due to evoked heart rate and blood pulse volume responses. Furthermore, there is the known coupling between heart rate and respiration, the latter response being not totally controlled by the *autonomic* nervous system.

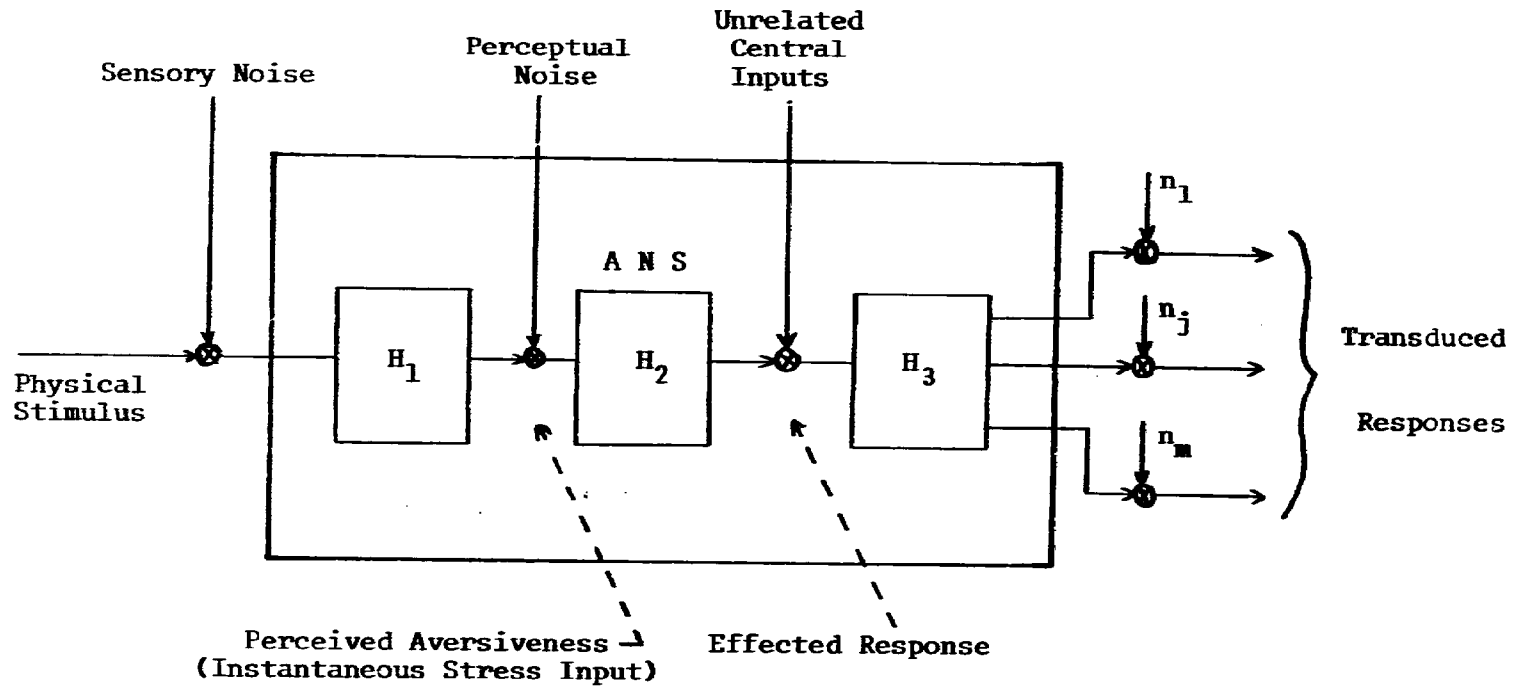


Figure 1. ANS Stress Input-Physiological Output Model

In summary, we hold no delusions about the simplicity of examining the global single block system, taking the outputs one at a time and treating the system as a single input-single output "black box" with additive output noise. Paired treatment of different evoked responses, however, may be expected to reveal valuable information about effector mechanisms and response transduction. It also may be stated at this point that there is no reason to expect the system to be linear or stationary. Careful design of the experiment, as shown in the following section, may alleviate these problems, however, and allow the utilization of quasi-linear and quasi-stationary system identification techniques.

2. EXPERIMENT

The principal objective in the design of the experiment described here is to overcome the apparent shortcomings of similar investigations conducted in a great many psychophysiological laboratories and to render the recorded signals amenable to stationary and linear identification techniques. Whereas in most similar experiments the subjects are presented with *"discrete"* sensory stimuli, such as bursts of noise, tones, flashes of light or electric shock, this is clearly unsatisfactory for the estimation of a describing function when the system is generally accepted as being non-linear and non-stationary. Common examples of systematic nonlinearities include perceptual threshold, psychophysical power functions and effector saturation. In general, wide-sense nonstationarities may be classified as either nonstationarities of the mean, generated by electrode, amplifier and membrane changes or tonic drifts, or of the mean square, due to peripheral sensory adaptation or more central habituation [2]. In the context of the sensory and perceptual noises of Fig. 1, our aim is to enhance SNR by minimizing these noises and to *"linearise"* the response by *driving* the system with a *continuous intense noxious stimulus*.

Most psychophysiological experiments tend to extract a restricted number of contrived measures which appear appropriate to a visual description of the data and are often not standardized among different laboratories. Examples of these are peak amplitudes, latencies, durations and rates of recovery of negative and positive components, time-to-peak amplitudes, basal level changes, and ratios of specific to non-specific discrete responses [3]. By introducing a *continuous* stimulus and recording a *continuous* response, we are therefore employing all available information and thus incorporating all of these measures.

Central to our objective is the ability to *"drive"* the ANS in a continuous, non-decrementing fashion, in order to eliminate the effects of habituation [4]. Although sinusoidal stimulation is a candidate waveform to accomplish this end, and, indeed, has been used by us in early studies [5], it has been nevertheless rejected due to problems of stationarity due to its predictability and the necessity of repeated presentation in order to cover a suitable representative range of frequencies and amplitudes [2, 6 (Ch. 9)]. For the purpose of presenting a dramatic example of driving, however, we present Fig. 2, where the bottom trace represents the time course of the envelope of a 0.1 Hz amplitude modulated (100 Hz) stimulus current (I_s) oscillating between 3.5 and 4.6 ma rms. Above this is the evoked palmar skin conductance response on the contralateral hand (C-SCR), followed by ipsi- and contra-lateral skin potential

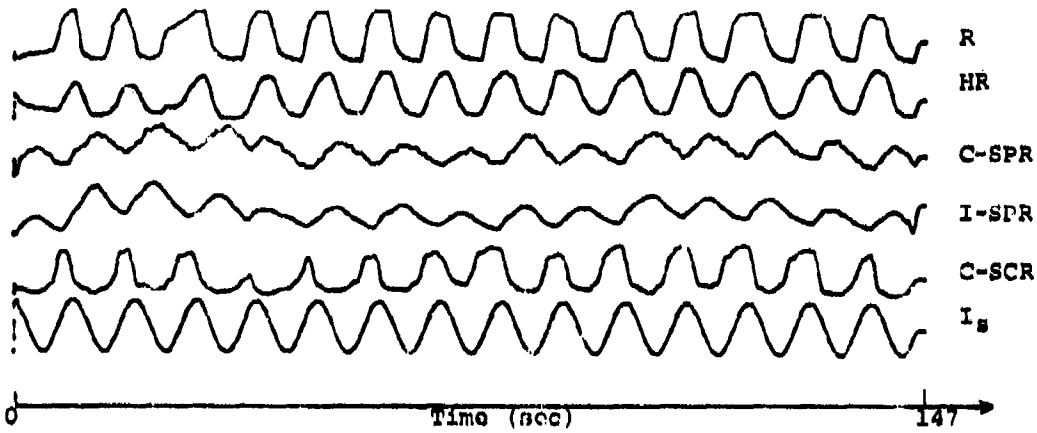


Figure 2. 0.1 Hz Sinusoidal Stimulation

responses (I-SPR and C-SPR respectively, both detrended), linearly interpolated heart rate (HR), and thermally monitored respiration (R). Clearly there is little or no monotonic response decrement over the 147 sec of data shown.

The use of a randomly presented forcing function must not be construed as solving all of the problems mentioned above, since linear differential equation parameters may very well change for different input spectra [6, Ch.9]. The advantages are many, however, among which are: a) mathematical tractability, allowing efficient utilization of the methods outlined in Part 3; b) forcing function "optimality", in that all modes of the response system may be "persistently excited" [7, 8 (Ch.11)]; c) "psychological stationarity", in that the subject is not able at any moment to predict the time course of the stimulus and is thus unlikely to habituate to it.

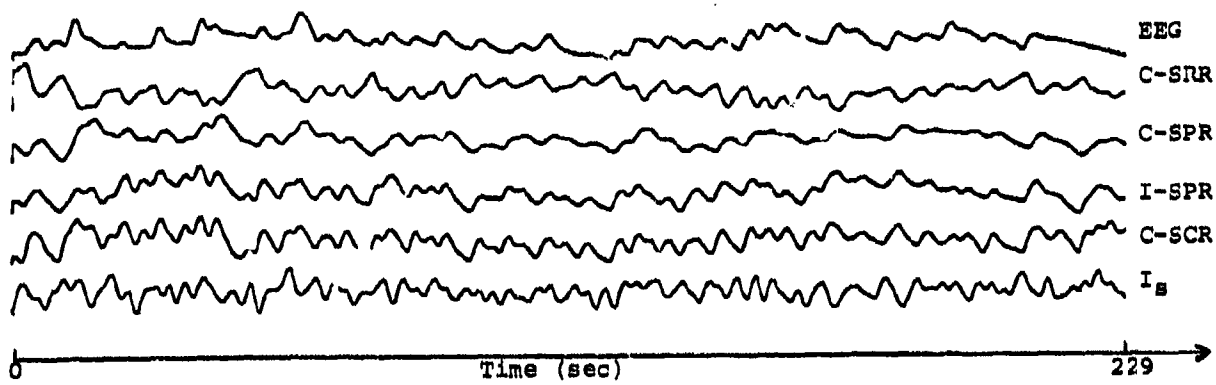


Figure 3. 0.3 Hz Low-pass Random Stimulation

Some sample data of a 0.3 Hz low-pass stimulus and its accompanying responses are given in Fig. 3, where from bottom to top are shown stimulus current envelope, contralateral SCR, ipsilateral SPR, contralateral SPR, contralateral SRR (skin resistance response), and DC coupled EEG (P_2-A_1). Each trace here represents 229 sec of data, and all responses, except SCR, have had polynomial trends removed for clarity. Clearly the responses evoked here are quite "labile"; however, we may no longer rely on visual evidence to show the existence of causality, or driving, but must conform to some stricter constraint. Although mathematically rigourous conditions for driving do exist [9], the sufficient condition that will be adopted here, given the unequivocal distinction between the driving and the driven measures, is that there exist a significant (lagged) input-output cross-correlation and/or range of non-zero squared coherence. Satisfaction of these criteria is illustrated in the next section.

3. ANALYSIS AND PRELIMINARY RESULTS

3.1 (Nonparametric) Spectral Analysis

Generally speaking, the philosophy of analysis pursued here follows closely that of Jenkins & Watts [10]. Presented with the sampled records of two continuous time physical processes, we commence by applying conventional *nonparametric* linear spectral analysis, as a flexible exploratory step in investigating the system's linearity and frequency structure. Our ultimate objective, however, is a *parsimonious "parametric"* representation of the system. We wish, in other words, to model the system as a set of linear differential equations, or corresponding linear difference equations, of as low an order as possible, and thence to estimate certain meaningful physical parameters, such as "damping factor" and "natural frequency" [11, 12].

All of the ensuing illustrations arise from the two time series traces shown at the bottom of Fig. 3, random stimulus current envelope (I_s) and contralateral skin conductance (SCR). All records have been analogically low-pass filtered at 8.0 Hz (-18 dB/octave) prior to digitization at a rate of 1 sample/56 msec. They are then zero-phase low-pass filtered at 0.45 Hz with a recursive IIR digital Butterworth filter [13], with leading and trailing points removed. The data are then hexadecimated, leaving 256 points, with an effective sampling rate of 1 sample/896 msec, representing 229.4 sec of data.

In Fig. 4 is shown the estimated cross correlation function relating I_s and SCR for positive and negative lags, computed by the sectioned FFT correlation algorithm of Rader [14], outlined in Chapter 11 of [15]. We see here clear evidence of a pure transport delay, which is estimated at 2.6 sec, a physiologically acceptable value which is comparable to those observed upon presentation of pulsed stimuli. The peak correlation estimate in Fig. 4 is ρ=0.6. However, it should be realized that the sample cross correlation function suffers from a very large correlation between neighbouring estimates and that it may be advisable first to prewhiten each series [10].

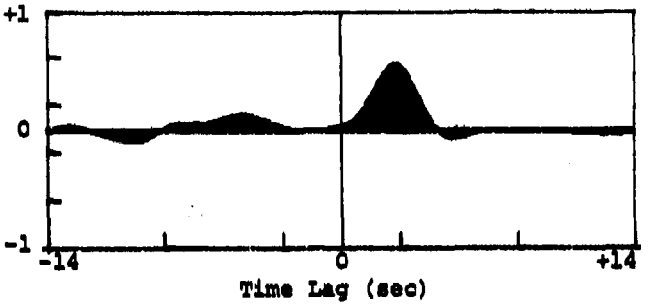


Figure 4. Estimated Cross-Correlation Function, I_a-SCR

A cross-spectral analysis of these same data is shown in Fig. 5. All spectra are computed by Welch's method of time averaging over short, modified overlapped periodograms [16, 17]. It is well known that application of such methods requires a great amount of judgement on the part of the analyst, in order to reduce the bias and variance of the spectral estimates while simultaneously maintaining adequate spectral resolution and reducing leakage due to finite length data windowing. In the method used here, a choice has been made of a suitable data window and data segment length, in a fashion analogous to the "window carpentry" and "window closing" of Jenkins & Watts [10]. For Fig. 5 all data segments were multiplied by a raised cosine, or Hann window. Not presented here, although quite illustrative of the amount of variance in the estimates, are the intermediate results for the same series as the data segmentation size is progressively decreased. Suffice it to say that the representation of Fig. 5 is very highly smoothed, and therefore highly biased. Since we do not expect any sharp spectral peaks in these data, we shall remain temporarily satisfied with these estimates; however, we shall shortly see how parametric spectral analysis may overcome many of the shortcomings of conventional spectral analysis. In any case, it is quite important to present the bottom graph of Fig. 5, a frequency domain representation of the equivalent convolution window resulting jointly from the data windowing and segment averaging procedures. The computed "equivalent noise bandwidth" (ENBW) is 0.0001 Hz and the number of "equivalent degrees of freedom" (EDF) for all estimates is 28.5 [18].

The next two curves are the power spectral estimates of the stimulus and response signals. The computed equivalent bandwidths (EBW) are 0.32 Hz and 0.21 Hz respectively, demonstrating obvious low-pass characteristics. Note that both graphs are plotted to a scale of 20 dB/division. We see here that the random stimulus has been designed to be "spectrally flat" beyond the expected bandwidth of the response signal, and thus to be effectively "white". (In fact, the stimulus current envelope was generated by a Hewlett-Packard 8057A Precision Noise Generator with the output low-pass filtered at 0.3 Hz. Since signals of different bandwidths have been found to be perceptually different, and since linear differential equation parameters may be expected to change for different input spectra, manipulation of the EBW of the stimulus current envelope is an experimental parameter which should be investigated further. A second experimental parameter, the variance of the stimulus envelope, is perhaps the most intangible, since there is no use expressing it in terms of physical units. In light of the discussion of Part 1, the level of stimulus

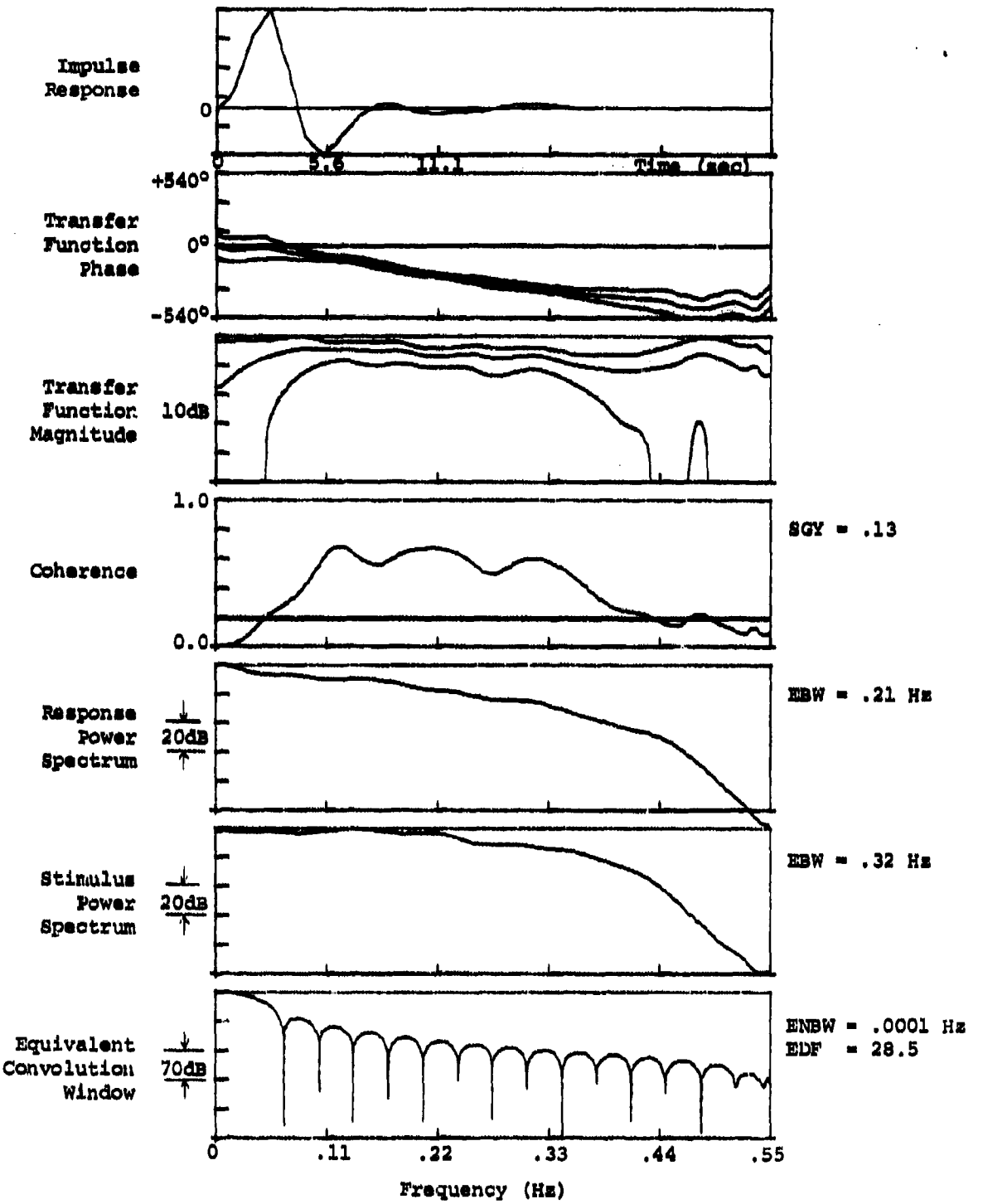


Figure 5.
Cross-Spectral Analysis: Stimulus Current — Contralateral Skin Conductance

aversiveness depends on the electrode preparation and on the subject's perception of the sensation. At present, we attempt to maintain all stimuli in a range between the subjective sensation levels "suprathreshold" and "intolerable".)

The sina qua non of the entire quasi-linear analysis is the graph of magnitude squared coherence. Here we observe a surprising degree of linear coherence, with maximum values of close to 0.7. The horizontal line across the bottom of the graph represents the value for which there is a 5% probability that the true coherence is zero [19, 20]. This threshold value, based on the number of EDF, is then used as the criterion for defining a *band of significant non-zero linear coherence*. The number labelled "SGY" represents an estimate of the *Shannon-Gelfand-Yaglom amount of information about the stimulus signal present in the response signal* [21, 22]. It is computed according to:

$$SGY = - \int \log (1-\gamma^2(f)) \, df$$

where $\gamma^2(f)$ is "linear thresholded coherence". The integration is thus performed for frequency (f) over the previously determined range of non-zero coherence. The SGY measure may then be considered as a unified or scalar measure of the degree of linearity relating the two signals.

The next two graphs are estimates of the linear transfer function magnitude and phase. Note that the magnitude graph has a scale of 10 dB/division, while the individual autospectra are plotted at 20 dB/division. The phase estimate has been "unwrapped", in order to produce a single smooth curve [23]. Both curves are bounded by 95% confidence limits, based on the coherence estimate and the EDF [10, 17]. It is clear that from the magnitude curve essentially no information may be derived, because of the low coherence at low and high frequencies. This is one of the shortcomings which we shall attempt to overcome in the following parametric analysis. The errors on the phase function appear quite small; however, this is due to the large vertical scale, and again there is a very loose bound at low and high frequencies. The two main pieces of information to be derived from this graph are: a) the apparently linear slope of approximately -2.5 to -3.0 seconds, in agreement with the previous cross-correlation estimate, thus confirming the causal stimulus-response relationship (i.e., driving); and b) an apparent positive phase intercept, demonstrating a systematic phase lead and thus implying that the subject is somehow *anticipating*, or *predicting*, the time course of the shock stimulus. When a 2.6 sec time lag is removed by "alignment" [10], this phase lead becomes much more evident, as the phase function estimate becomes essentially horizontal and positive. (In addition, the coherence estimate becomes much improved by the operation of aligning, as predicted by Jenkins & Watts [10], reaching values of 0.8.)

Finally, the top graph in Fig. 5 is an estimate of the system's linear impulse response in the time domain. It is derived by inverse Fourier transformation from the transfer function magnitude and phase estimates. Note the peak value at 2.6 sec. Once again, a more representative (less biased) estimate of the impulse response is obtained when the 2.6 sec lag is removed by aligning the two series. Bearing in mind that discrete Fourier analysis of finite length data records is a finite approximation to an infinite process, this impulse response estimate is *non-parametric*, since an effectively infinite number of parameters is required to specify the process. Specifically, the number of parameters equals the number of data samples (256). Clearly, the

series may be truncated at some point with very little error; however, there will still be many more parameters than needed for a suitably chosen parametric model, as shown in the following section.

3.2 Autoregressive Moving Average Models

In this section there will be presented two approaches for fitting a finite order parametric series to the data of the previous section, only one of which will be pursued. Both are able to overcome the necessity of exercising arbitrary judgement in windowing and sectioning in conventional spectral analysis, since the model orders may be "objectively" determined by an information theoretic criterion [24, 25]. In addition, whereas conventional spectral analysis methods tend to underestimate sharp spectral peaks and overestimate spectral troughs, the present procedures tend to overcome these problems while maintaining asymptotically efficient statistical estimation properties. On the other hand, the concepts of degrees of freedom and bandwidth in conventional spectral analysis do not precisely apply in parametric modelled spectral analysis, and, in addition, some of the flexibility of conventional spectral analysis is lost.

In the discussion of cross-spectral analysis no assumptions were made as to whether or not one series was causal to the other. This in fact could not have been assumed a priori, since this is precisely what was sought when we endeavoured to prove that the stimulus was driving the response, by means of identifying a significant lagged cross-correlation coefficient and a band of significant non-zero spectral coherence. Our tacit assumption, therefore, was that the two signals had arisen "on a similar footing" [10], and we proceeded to show causality. In other words, an input-output causal system is a special case of a (bivariate) system of two signals which are mutually causal.

In the general case, we therefore assume a system where the two signals of interest, $x_1(t)$ and $x_2(t)$, are driven by two (possibly auto-correlated) noise sources. Such a scheme is illustrated in Fig. 6(a). Here $\xi_1(t)$ and $\xi_2(t)$ are uncorrelated white noise processes (with variances σ_1^2 and σ_2^2). The two filters $G_{11}(z)$ and $G_{22}(z)$ account for any autocorrelation in the noise sources $e_1(t)$ and $e_2(t)$.† The filters $H_{12}(z)$ and $H_{21}(z)$ account for any coupling between the two signals $x_1(t)$ and $x_2(t)$. The structure of Fig. 6(a) may be expressed more tractably in matrix form as:

$$\sum_{k=0}^{p} H_k \cdot \underline{x}(t-k) = \sum_{k=0}^{q} G_k \cdot \underline{\xi}(t-k) \qquad (1)$$

where $\underline{x}(t)=[x_1(t)\ x_2(t)]'$ and $\underline{\xi}(t)=[\xi_1(t)\ \xi_2(t)]'$ are sampled 2-vectors,

†All transfer functions are given here in terms of the z-transform, where the z-transform $X(z)$ of a sampled sequence $x(n)$ is defined as $X(z)=\sum_{n=-\infty}^{\infty} x(n)z^{-n}$, where z is complex. Setting $z=e^{j\omega}$, or in other words projecting $X(z)$ onto the unit circle $|z|=1$, we arrive at the equivalent Fourier transform values, $X(e^{j\omega})$. For a finite duration sequence $x(n)$, the discrete Fourier transform (DFT) corresponds to equally spaced samples on the unit circle of its z-transform [15].

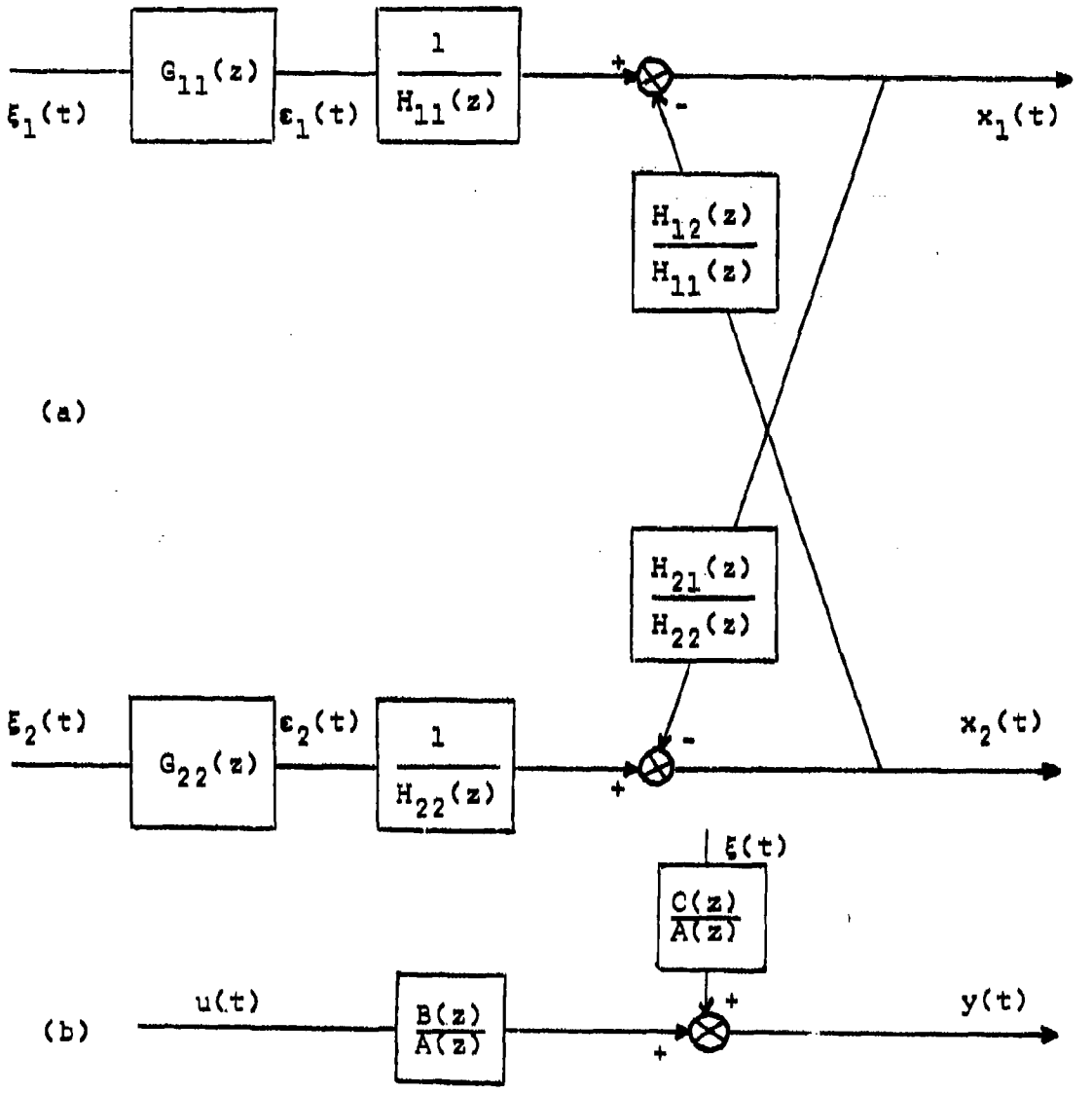


Figure 6. Generalized Autoregressive Moving Average (ARMA) Representation
(a) Bivariate linear process
(b) Linear input-output model

and H_k and G_k are 2 x 2 matrices,

where $H_0=G_0=I$, the identity matrix,

and $G_{12}=G_{21}=0$, for all k (i.e., G is a diagonal matrix).

In other words, we are assuming a finite order structural model where we express the bivariate observations of $\underline{x}(t)$ as a linear combination of a finite number of past observations plus a finite linear combination of past and present random unpredictable inputs:

$$\underline{x}(t) = -H_1\underline{x}(t-1)-H_2\underline{x}(t-2)-\ldots-H_p\underline{x}(t-p) + \underline{\xi}(t)+G_1\underline{\xi}(t-1)+\ldots+G_q\underline{\xi}(t-q) \qquad (2)$$

This form may of course be generalized beyond 2 x 2 matrices for a multivariate system.

For the special case of q=0, we have a pure *autoregressive* (AR) model, where the inputs $\underline{\varepsilon}(t)=\underline{\xi}(t)$ are already not autocorrelated. When the order of this estimated scheme, p, is the same as the order of the model that generated the data, the estimated parameters are then maximum likelihood estimates. In addition, the AR spectral estimates tend to be unbiased, and their variances asymptotically approach those of conventional spectral analysis. The spectral resolution of the AR estimator is superior to that of windowed periodogram spectral analysis; however, it is highly sensitive to the signal-to-noise ratio [27]. Least squares estimation of the parameters of the AR model reduces to the solution of a set of 2pth order linear equations based on estimates of the covariance matrix, and may be solved recursively in a straightforward manner [28, 29, 30].

If the assumption q=0 is not true of the real underlying process, however, estimates of the system parameters H will tend to be biased [7, 8, 12]. We must then introduce the G parameters to account for any autocorrelation in $\underline{\varepsilon}(t)$. Since we are expressing:

$$\underline{\varepsilon}(z) = G(z)\cdot\underline{\xi}(z), \quad \text{or} \quad \underline{\varepsilon}(t) = \sum_{k=0}^{q} G_k\cdot\underline{\xi}(t-k),$$

G(z) is a *moving average* (MA) filter, and the model for q>0 now becomes an *autoregressive moving average* (ARMA) representation. In this case we may, without loss of generality, set p=q. Least squares estimation of the H and G parameters for multivariate series is no longer a linear problem, but it may be variously solved by such methods as the *two stage least squares* (2SLS) method developed by Gersch & Yonemoto [26].

For both the AR and ARMA models, it is possible to select p, the order of the model, from among a choice of different model orders. The usual and very versatile method used is *Akaike's information criterion* (AIC), which prescribes the model order k for which the value AIC(k):

$$AIC(k) = N\cdot\ln|S(k)| + 2d(1+ds), \quad k=0,1,2.... \qquad (3)$$

is a minimum. In equation (3) N is the number of data points; $|S(k)|$ is the determinant of the residual sum of squares and cross product matrices (corresponding to a regression analysis interpretation of model fitting); d is the

model dimension (d=2 for the bivariate scheme of Fig. 6(a)); and s is the number of matrices fitted (s=p for the AR model and s=p+q=2p for the ARMA model) [24, 25, 26].

The discussion of Fig. 6(a) has served as a framework for understanding how to model two time series arising on an equal footing from the finite order structural model described by equation (1). Although the parameters of equation (1) could have been estimated by the 2SLS method mentioned [26], we have chosen at this point not to present these estimates, but rather to simplify the model by using our knowledge of the underlying causality in the system. In Part 3.1 we have seen, from the cross-correlation lag (or phase function slope) and band of non-zero coherence, that the input simulus is *driving* the recorded response. We have also speculated in Part 1 that there is essentially no feedback component whereby the response might influence the stimulus. We therefore see no need to fit both the H_{12} and the H_{21} parameters of Fig. 6(a) and equation (1). In theory one of them should turn out zero; in practice, due to computation errors and errors inherent in estimating sample covariances, this is not so.

We therefore set $x_1(t)=u(t)$, $x_2(t)=y(t)$, $\xi_2(t)=\xi(t)$, and $H_{12}=0$, $H_{22}=A$ $H_{21}=-B$, $G_{22}=C$, for all k. Equation (1) may now be written:

$$\sum_{k=0}^{p} A_k y(t-k) = \sum_{k=0}^{p} B_k u(t-k) + \sum_{k=0}^{p} C_k \xi(t-k) \qquad (4)$$

where A_k, B_k, C_k are scalars and u(t) and y(t) represent the system's sampled input and output respectively. Without loss of generality, we also set $A_0=C_0=1$, and we also set $B_0=0$. This new model, which is actually a special case of that of Fig. 6(a), is shown schematically in Fig. 6(b). Conceptually, we may look upon this representation as a familiar single input single output system with additive observation noise, where we relate the variables to the physical measures shown by the single block representation of Fig. 1.

Least squares estimation of the parameters of the generalized linear difference equation (4) may be carried out with various documented techniques [8, 11, 12]. We have used a 2SLS method similar to that described previously for the estimation of multivariate ARMA model parameters [31]. Briefly, this is a non-iterative computational procedure that is suitable for fitting large order parametric ARMA models to data, given the sample covariance estimates $R_{uu}(k)$, $R_{uy}(k)$, $R_{yu}(k)$, $R_{yy}(k)$. The first stage computation assumes an uncorrelated residual $\xi(t)$, and fits a pure AR model ($C_k=0$, $k>0$), by solving a set of linear least squares equations of order 2p, by block Toeplitz matrix inversion. These estimates are then used to estimate the covariances $R_{u\xi}(k)$, $R_{\xi u}(k)$, $R_{y\xi}(k)$, $R_{\xi y}(k)$, and $R_{\xi\xi}(k)$. The second least squares stage then inverts a 3p block Toeplitz matrix to provide estimates of the 3p A, B and C parameters. The AIC is computed for successive estimations at each stage for increasing order p, until an optimum order p is found from the minimum AIC(p).

In Fig. 7 are shown the results of applying the 2SLS input-output algorithm to the I$_s$-SCR data of Fig. 3. The graphs are drawn to the same scale as those generated by conventional spectral analysis in Fig. 5, for purposes of comparison. No longer shown are the equivalent convolution window, since it has no meaning here, and the coherence function, which can not be estimated by this method. The reason for this is that the coherence function is a measure of the cross power spectrum magnitude, which is normalized by the auto-spectra of the two individual time series, to produce values between 0 and 1. Since the input and output auto-spectral estimates are not produced by the input-output ARMA method, they are not available for normalization of coherence. In other words, we have eliminated estimates of H_{11} and G_{11} in going from Fig. 6(a) to 6(b), or from equation (1) to (4). Coherence estimates are therefore a product of bivariate (or multivariate) modelling, where the series are assumed to arise on an equal footing. Coherence could have been estimated from the model of equation (1); however, we have not as yet implemented an estimation procedure for the H and G parameters, and so the coherence estimate is absent in Fig. 7.

The stimulus and response auto-power spectra in Fig. 7 have each been estimated as pure autoregressive scalar processes [32, 33]. Each has been fitted to p=7th order linear models, from the minimum AIC. They are presented here for comparison with the corresponding graphs of Fig. 5. Note that the EBW's are very close, and that, within this bandwidth, the shapes are also very similar.

The results of the 2SLS programme suggested a model of order p=5, and the top three graphs of Fig. 7 are produced by these estimates. Most valuable is the transfer function magnitude which, drawn to the same scale as that of Fig. 5, clears up much of the ambiguity of that figure. The phase and impulse response curves also match up very well.

From the transfer function magnitude curve we see that it is not unreasonable to assume that we are dealing with an analogue of a structural system of damped second order coupled sub-systems as mentioned briefly at the beginning of Part 3.1. Since the system has been excited by a random forcing function, we are able, from the estimated parameters of the linear *difference equation* (4), describing a mixed ARMA discrete time series, to estimate the damping coefficients and natural frequencies of a corresponding n degree of freedom second order coupled *continuous differential equation model* [34, 35, 36, 37]. These values are specified by the poles of the transfer function, as determined by estimates of the AR parameters (A values of equation (4)). In addition, estimates of the coefficients of variation of the structural system parameter estimates can be derived from the ARMA parameters (A's and B's) [35, 38]. Accordingly, for the p=5 model identified from Fig. 7, we present the following values:

Pole Locations	Damping Factor	Natural Frequency
0.44 ± 0.51j	0.42	0.17 Hz
-0.23 ± 0.63j	0.32	0.23 Hz

Note that pole locations are given in complex z-plane coordinates. The coefficients of variation have not yet been computed at this point.

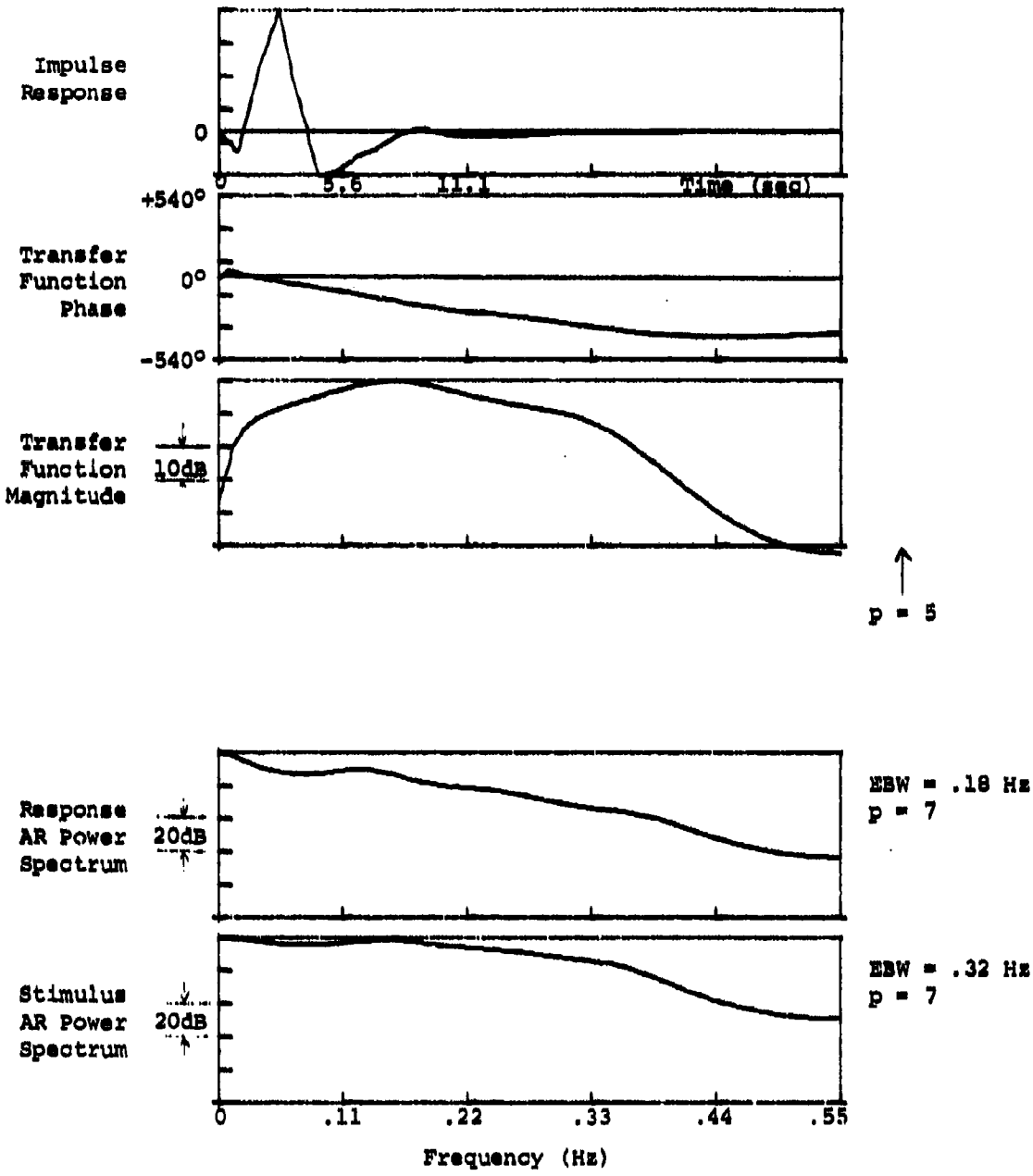


Figure 7. 2SLS ARMA fit for I_s-SCR data (p=5)

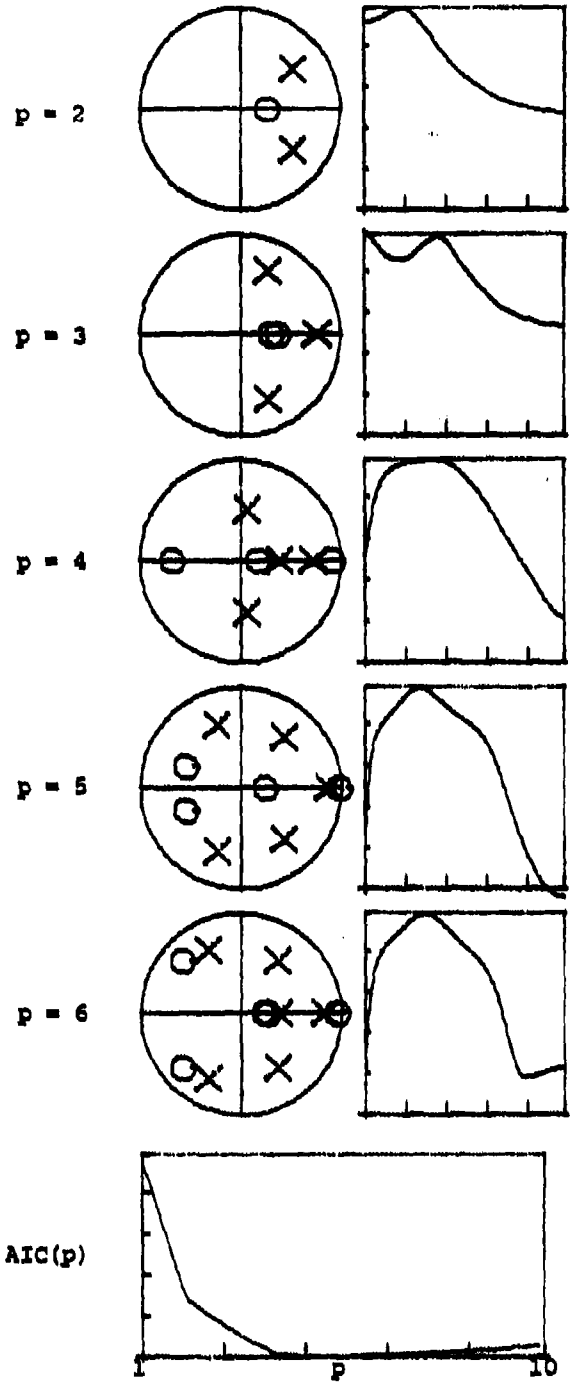


Figure 8.
2SLS ARMA fit to I_{sc}-SCR data: Complex pole(X)-zero(O) configuration (large circles are $|z|=1$) and corresponding linear transfer functions for increasing model order, p. All transfer functions have 10dB/division ordinates and 0 to 0.55 Hz abscissae. Bottom graph shows corresponding AIC for p=1,...,10.

We conclude by examining the modelled process and speculating upon whether it is possible to simplify it somewhat. In the ideal case, if the underlying process is a true finite order ARMA process, our method of evaluation of the AIC should approach the correct system order. Given that we have not complete linearity and that we expect many extraneous noises, we expect that the residual error will be quite large and that our procedure will possibly *overfit* the model. With this in mind, we present Fig. 8, which is a representation of the iterative results of the 2SLS ARMA identification procedure for the same data, for successive orders of p from 2 to 6. Recall that the results of Fig. 7 are for p=5. The left-hand column of graphs are z-plane estimates of the pole (X) and zero (O) locations. Each large circle represents the curve $|z|=1$. In the right-hand column of graphs are the transfer function magnitude estimates corresponding to the pole-zero plots. All vertical scales are set equal at 10 dB/division. At the bottom of Fig. 8 is a graph of AIC(p) vs p for a set of successively computed model orders. We see that, as the model order is increased for p>2, the poles and zeroes shift around somewhat, but that there is a great tendency for the additional X's and O's to cancel each other. We should therefore seriously consider the conclusion that p=2 is of sufficient order to model our data, and that the underlying input-output system may be modelled by a single damped second order system [39]. The poles, in this case, are at $0.53 \pm 0.40j$, and we have a damping coefficient of 0.54 and a natural frequency of 0.14 Hz.

4. CONCLUSION

In this paper we have presented a unique psychophysiological experiment for probing the dynamic response characteristics of the human autonomic nervous system. By presenting subjects with a localized, continuously fluctuating, low-pass, randomly amplitude modulated, noxious electric current, and by recording a number of elicited responses, we have shown that it is possible to drive the responses in a non-decrementing fashion. We then attempted to estimate, to a first approximation, some of the properties of an input-output describing function. With no assumptions about the underlying system structure, we present the following list of potentially meaningful identifiers, or characteristic system parameters.

1. estimated cross-correlation coefficient
2. estimated transport delay
3. estimated response bandwidth
4. width of identified band of estimated non-zero coherence
5. Shannon-Gelfand-Yaglom measure of amount of information about the input in the output
6. zero frequency phase lead (or lag)

We then proceeded to speculate that the underlying generating mechanism is a one degree of freedom, damped second order system, whose identified damping coefficient and natural frequency may also be taken as meaningful system identifiers.

Two approaches of system identification were used. The first, non-parametric, method utilized "conventional" windowed periodogram cross-spectral analysis. Taken in terms of input-output models, this effectively reduces to an infinite order moving average (MA) representation. Having mentioned many of the advantages and disadvantages of this approach, we proceeded to model the system with finite order linear difference equations, as an autoregressive moving average (ARMA) process. The implications of a bivariate time series approach were discussed, but the sub-case of a causally defined input-ouput system was pursued. Sample results for stimulus current input—skin conductance response output were produced by the two approaches and were compared. We wish to emphasize the flexibility and relative objectiveness of the time series approach, in conjunction with the application of Akaike's information criterion (AIC). Although not yet in widespread use in the area of human operator manual control modelling, we foresee an increase in its utilization [40].

Some potential applications of the experiment presented here are:

a) the study of man's ability to tolerate various intensities and types of stress, for purposes of personnel selection or training;
b) the study of the effect of summation of individual stressors;
c) the examination of the possible effects of "autonomic resonance" induced by cyclically presented arousal stimuli;
d) the identification of "autonomic lability" of various population groups that may be especially susceptible to certain psychiatric disorders [2, 41].

We are presently, in our laboratory, commencing an investigation to gather data from a group of psychiatric outpatients.

ACKNOWLEDGEMENT

The authors would like to express our great appreciation to Prof. Will Gersch of the Department of Information and Computer Sciences of the University of Hawaii for use of his 2SLS system identification programme and for his invaluable contribution of insight, suggestions and answers to our questions.

REFERENCES

1. DV Cross, B Tursky, & M Lodge, "The Role of Regression and Range Effects in Determination of the Power Function for Electric Shock," *Perception & Psychophysics*, Vol. 18, 1975, 9-14.

2. P Milgram & J Senders, "A Spectral Characterization of Human Autonomic Responsivity," *Proc. Intn'l Conf. Cybernetics & Society, Tokyo, Nov. 1978*, 187-190.

3. PH Venables & MJ Christie, "Mechanisms, Instrumentation, Recording Techniques, and Quantification of Responses," in WF Prokasy & DC Raskin (ed's), *Electrodermal Activity in Psychological Research*. New York: Academic Press, 1973, Ch. 1, 1-124.

4. FK Graham, "Habituation and Dishabituation of Responses Innervated by the Autonomic Nervous System," in HVS Peeko & MJ Herz (ed's), *Habituation*. New York: Academic Press, 1973, Ch. 5, 163-218.

5. VL Senders, JW Senders & B Tursky, "An Investigation of Continuously Modulated Skin Potential Responses," *Annual Meeting of Psychonomic Society*, 1972.

6. TB Sheridan & WR Ferrell, *Man-Machine Systems: Information, Control, and Decision Models of Human Performance*. Cambridge, MA: MIT Press, 1974.

7. KJ Astrom & P Eykhoff, "System Identification—A Survey," *Automatica*, Vol. 7, 1971, 123-162.

8. GEP Box & GM Jenkins, *Time Series Analysis: Forecasting and Control*. Holden-Day, (Revised Edition) 1976.

9. W Gersch, "Causality or Driving in Electrophysiological Signal Analysis," *Mathematical Biosciences*, Vol. 14, 1972, 177-196.

10. GM Jenkins & DG Watts, *Spectral Analysis and Its Applications*. Holden-Day, 1968.

11. TC Hsia, *System Identification: Least Squares Methods*. Lexington, MA: Lexington Books, 1977.

12. P Eykhoff, *System Identification: Parameter and State Estimation*. Wiley, 1974.

13. SD Stearns, *Digital Signal Analysis*. Hayden Book Co., 1975.

14. CM Rader, "An Improved Algorithm for High-Speed Autocorrelation with Applications to Spectral Estimation," *IEEE Trans Audio Electroacoustics*, Vol. AU-18, Dec. 1970, 439-441.

15. AV Oppenheim & RW Schafer, *Digital Signal Processing*. New Jersey: Prentice-Hall, 1975.

16. PD Welch, "The Use of Fast Fourier Transform for the Estimation of Power Spectra: A Method Based on Time Averaging Over Short, Modified Periodograms," *IEEE Trans Audio Electroacoustics*, Vol. AU-15, June 1967, 70-73.

17. JS Bendat & AG Piersol, *Random Data: Analysis and Measurement Procedures*. Wiley-Interscience, 1971.

18. FJ Harris, "On the Use of Windows for Harmonic Analysis with the Discrete Fourier Transform," *Proc. IEEE*, Vol. 66, Jan. 1978, 51-83.

19. GC Carter, "Receiver Operating Characteristics for a Linearly Thresholded Coherence Estimation Detector," *IEEE Trans Acoustics, Speech & Signal Processing*, Vol. ASSP-25, Feb. 1977, 90-92.

20. RH Jones, "Spectrum Estimation and Time Series Analysis—A Review," *Intn'l Symposium on Probability & Statistics in the Atmospheric Sciences*, June 1971.

21. W Gersch & BR Tharp, "Spectral Regression—Amount of Information Analysis of Seizures in Humans," in P Kellaway & I Petersen (ed's), *Quantitative Analytic Studies in Epilepsy*. New York: Raven Press, 1976, 509-532.

22. W Gersch, J Yonemoto, P Naitoh, "Automatic Classification of Multivariate EEGs Using an Amount of Information Measure and the Eigenvalues of Parametric Time Series Model Features," *Computers & Biomed Research*, Vol. 10, 1977, 297-318.

23. JM Tribolet, "A New Phase Unwrapping Algorithm," *IEEE Trans Acoustics, Speech & Signal Processing*, Vol. ASSP-25, April 1977, 170-177.

24. H Akaike, "A New Look at the Statistical Model Identification," *IEEE Trans Automatic Control*, Vol. AC-19, Dec. 1974, 716-723.

25. H Akaike, "Canonical Correlation Analysis of Time Series and the Use of an Information Criterion," in RK Mehra & DG Lainiotis (ed's), *System Identification (Advances and Case Studies)*. New York: Academic Press, 1976, 27-96.

26. W Gersch & J Yonemoto, "Parametric Time Series Models for Multivariate EEG Analysis," *Computers & Biomed Research*, Vol. 10, April 1977, 113-125.

27. M Kaveh & GR Cooper, "An Empirical Investigation of the Properties of the Autoregressive Spectral Estimator," *IEEE Trans Information Theory*, Vol. IT-22, May 1976, 313-323.

28. P Whittle, "On the Fitting of Multivariate Autoregressions, and the Approximate Canonical Factorization of a Spectral Density Matrix," *Biometrika*, Vol. 50, 1963, 129-134.

29. RH Jones, "Identification and Autoregressive Spectrum Estimation," *IEEE Trans Automatic Control*, Vol. AC-19, Dec. 1974, 894-897.

30. W Gersch, "Spectral Analysis of EEG's by Autoregressive Decomposition of Time Series," *Mathematical Biosciences*, Vol. 7, 1970, 205-222.

31. W Gersch, S Braun, FJ Martinelli, J Yonemoto, "A Two-Stage Least Squares Method for the Identification of Input-Output Vibration Data Systems," *Symposium on Applications of Computer Methods in Engineering, U of S California*, Aug. 1977.

32. J Makhoul, "Linear Prediction: A Tutorial Review," *Proc IEEE*, Vol. 63, April 1975, 561-580.

33. DG Childers (ed), *Modern Spectrum Analysis*. New York: IEEE Press, 1978.

34. W Gersch & S Luo, "Discrete Time Series Synthesis of Randomly Excited Structural System Response," *J Acoustical Society America*, Vol. 51, 1972, 402-408.

35. W Gersch & DA Foutch, "Least Squares Estimates of Structural System Parameters Using Covariance Function Data," *IEEE Trans Automatic Control*, Vol. AC-19, Dec. 1974, 898-903.

36. W Gersch, GT Taoka, R Liu, "Structural System Parameter Estimation by Two-Stage Least-Squares Method," *Proc ASCE*, Vol. 102, No. EM5, Oct. 1976, 883-899.

37. W Gersch & R Liu, "Time Series Methods for the Synthesis of Random Vibration Systems," *Trans ASME*, No. 76-APM-A, 1976.

38. W Gersch, "On the Achievable Accuracy of Structural System Parameter Estimates," *J Sound & Vibration*, Vol. 34, May 1976, 63-79.

39. AJW van den Boom & AWM van den Enden, "The Determination of the Order of Process- and Noise Dynamics," in P Eykhoff (ed), *Identification and System Parameter Estimation. Proc. 3rd IFAC Symposium, The Hague/Delft, The Netherlands, June 1973.* North-Holland Publishing Co., 1973, TS-2, 929-938.

40. SM Shinners, "Modeling of Human Operator Performance Utilizing Time Series Analysis," *IEEE Trans Systems, Man & Cybernetics*, Vol. SMC-4, Sept. 1974, 446-458.

41. AF Ax, "Goals and Methods of Psychophysiology," *Psychophysiology*, Vol. 1, July 1964, 8-25.

A MODEL FOR MANUAL DECELERATING APPROACHES TO HOVER

Robert K. Heffley
Senior Research Engineer
Systems Technology, Inc.
2672 Bayshore-Frontage Road
Mountain View, California 94043

ABSTRACT

A mathematical model is presented which describes the manual deceleration of a helicopter or VTOL aircraft from high speed flight to hover. The model combines the rules of visual perception with the crossover model of the human operator. The result is a two parameter, time varying description of closure speed versus range for which time and deceleration can be solved directly. The model compares well with flight test measurements of helicopter deceleration maneuvers. One potential use of the model is as a simulator validation tool by comparing simulator-measured model parameters with in-flight measurements. Extension of the model to vertical and lateral axes is briefly discussed.

LIST OF SYMBOLS

A	Effective size of viewed object
c	Constant (eq. 6)
h	Altitude
$\dot{h}$	Altitude rate
k	Constant of proportionality for range
k_h	Constant of proportionality for height
k_y	Constant of proportionality for lateral offset
n	Exponential constant (Eq. 6)
R	Range from hover point
$\dot{R}$	Range rate
$\ddot{R}$	Deceleration
R_p	Perceived range
$\ddot{R}_{pk}$	Peak deceleration
y	Lateral offset
$\dot{y}$	Lateral velocity
t	Time

(Numerical subscripts are used to denote corresponding range-time combinations.)

INTRODUCTION

The manually controlled decelerating approach to hover in a helicopter or VTOL aircraft is normally viewed as a time varying maneuver for which conventional analysis techniques do not apply. A mathematical model has been formulated, however, which combines the crossover model with the effects of visual perception and yields a simple guidance law which agrees well with in-flight measurements of pilot actions.* Although the model is time varying, it permits closed form solutions for speed, acceleration, and time as functions of range. In addition, the same ideas applied to the deceleration model can also be extended to vertical and lateral flight path guidance.

In the following pages we describe the basis of the model, its practical formulation, agreement with flight data, and potential applications.

THEORETICAL BASIS

The hypothesis used to formulate the decelerating approach model is that range rate is varied in direct proportion to perceived range, i.e.,

$$\dot{R} = -k R_p \tag{1}$$

where $\dot{R}$ is actual range rate

R_p is perceived range

k is the constant of proportionality

This carries the implication of rate-command-like behavior implicit in the crossover model of the human operator as described in Ref. 1 with allowance for visual perspective effects.

The key to describing the visual aspect is provided in Ref. 2 where the psychological measurements of apparent size and distance are related to various analytically derived relationships. It is shown that perceived range, R_p, is related to true range, R, by:

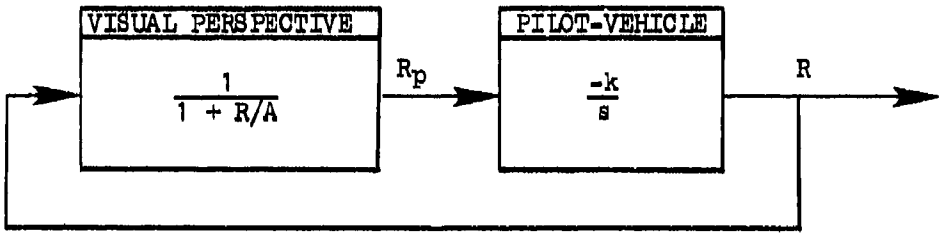
$$R_p = \frac{R}{1 + R/A} \tag{2}$$

where the length A is a measure of the effective size of the object being viewed.

Thus the psychological perception of range combined with the crossover model for human operator behavior leads to the following block diagram description for control of range-to-go:

---

* The model was developed in direct support of Naval Air Engineering Center Contract N68335-78-C-2019 and Naval Air Development Contract N62269-77-C-0509.



This can be written as:

$$\dot{R} = -k \frac{R}{1 + R/A} \tag{3}$$

SOLUTION OF THE GUIDANCE LAW

The above guidance law yields a direct closed-form solution if the crossover gain, k, and the effective size parameter, A, are assumed constant. Starting with Eq. 3,

$$\frac{dR}{dt} = \frac{-k\,R}{1 + R/A}$$

and integrating

$$\int_{t_1}^{t_2} dt = -\frac{1}{k} \int_{R_1}^{R_2} \left(\frac{1}{R} + \frac{1}{A}\right) dR$$

we obtain the result

$$t_2 - t_1 = \frac{1}{k}\left(\ln \frac{R_1}{R_2} + \frac{R_1 - R_2}{A}\right) \tag{4}$$

where R_1 is range at time t_1

and R_2 is range at time t_2

Inertial deceleration with range, $\ddot{R}$, can be computed by differentiating Eq. 3, i.e.,

$$\ddot{R} = \frac{k^2 R}{(1 + R/A)^3} \tag{5}$$

A maximum deceleration can thus be found to occur at a range equal to A/2 with a magnitude of $4k^2A/27$. Plots of range rate and deceleration therefore have the characteristic shapes shown in Fig. 1.

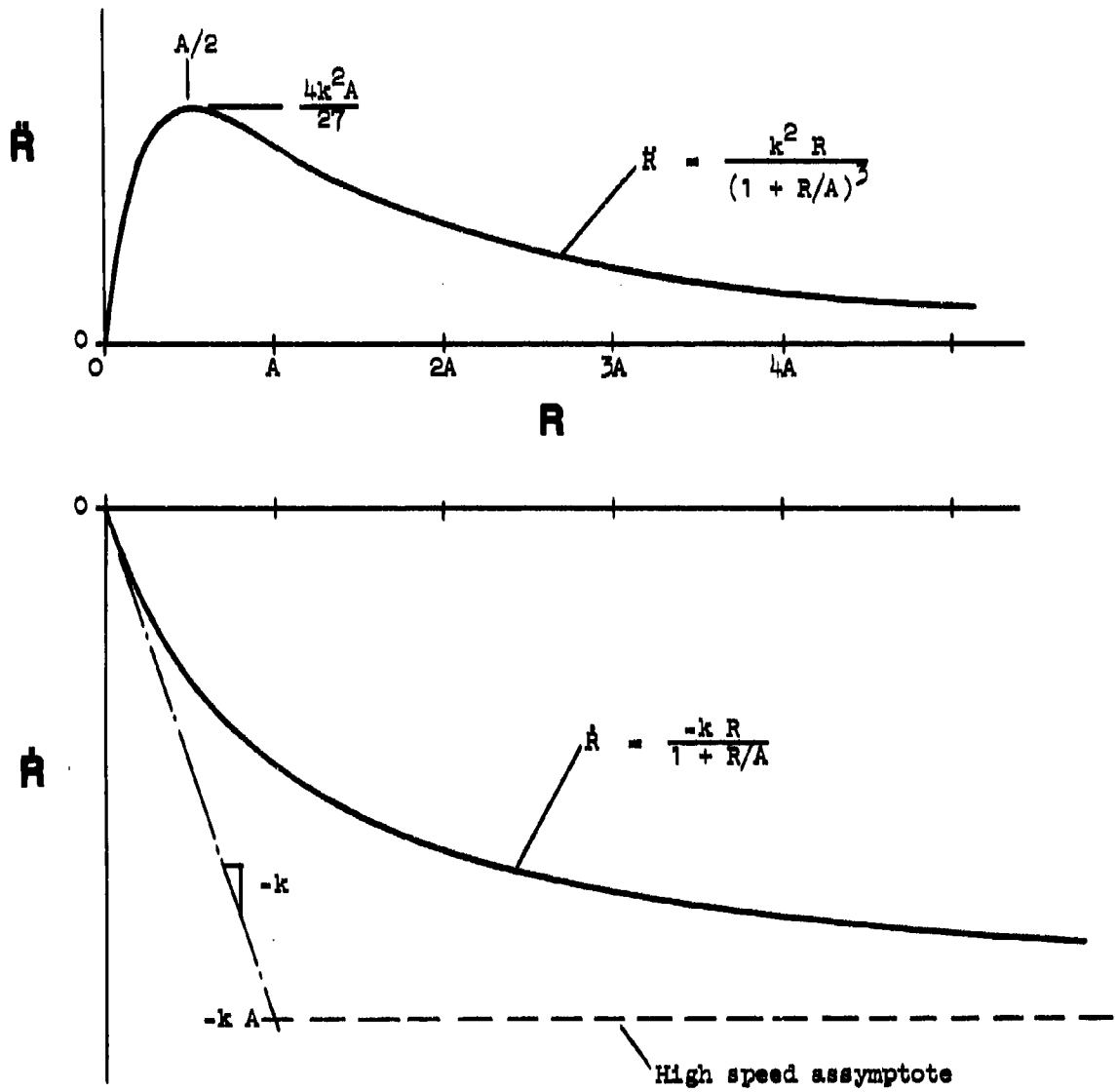


Figure 1. Deceleration and Closure Speed Versus Range

AGREEMENT WITH FLIGHT DATA

The analytical model thus described reflects the essential features of deceleration profiles obtained from flight test measurements. Reference 3 contains data based on two hundred approaches by various pilots using four types of helicopters. The approaches were started at combinations of three airspeeds and three altitudes. Representative initial conditions for the approach were considered to be 80 kt airspeed and 1,000 ft altitude. A typical deceleration profile taken directly from Ref. 3 is shown in Fig. 2 with two analytical guidance model solutions superimposed, assuming k and A to be constants. The model parameters k and A are taken to be 0.25/sec and 600 ft in one case and 0.30/sec and 400 ft in the other. Note that the characteristic shape of the deceleration profile is followed using either pair of values for k and A although one pair matches the longer ranges better and the other pair, the shorter ranges. This may reflect a shift in the effective field of interest by the pilot between long range and short range. At longer ranges the field of interest may encompass the overall landing area, hence a larger value of A; and at shorter ranges the pilot may focus only on the precise landing spot with a correspondingly smaller A. Nevertheless, a reasonably accurate deceleration profile is given by a single set of (constant) model parameters.

The method used to pick a value for A can be based on an empirical relationship noted in Ref. 3, that is,

$$\frac{\dot{R}^2}{\ddot{R}} = c\,R^n \tag{6}$$

where c and n are constants

The analytical guidance model, at the same time, can be represented in the following form:

$$\frac{\dot{R}^2}{\ddot{R}} = R(1 + R/A) \tag{7}$$

Therefore, to solve for A we can match $\dot{R}^2/\ddot{R}$ over the range of the deceleration maneuver. The results of this match for A = 600 ft are shown in Fig. 3.

The second parameter, k, can be chosen by matching peak deceleration. Thus with the A obtained using Eq. 7 we can solve for k using the theoretical peak $\ddot{R}$:

$$k = \sqrt{\frac{27\,\ddot{R}_{pk}}{4A}}$$

Therefore, if $\ddot{R}_{pk}$ is 0.15 g, then k is 0.25/sec.

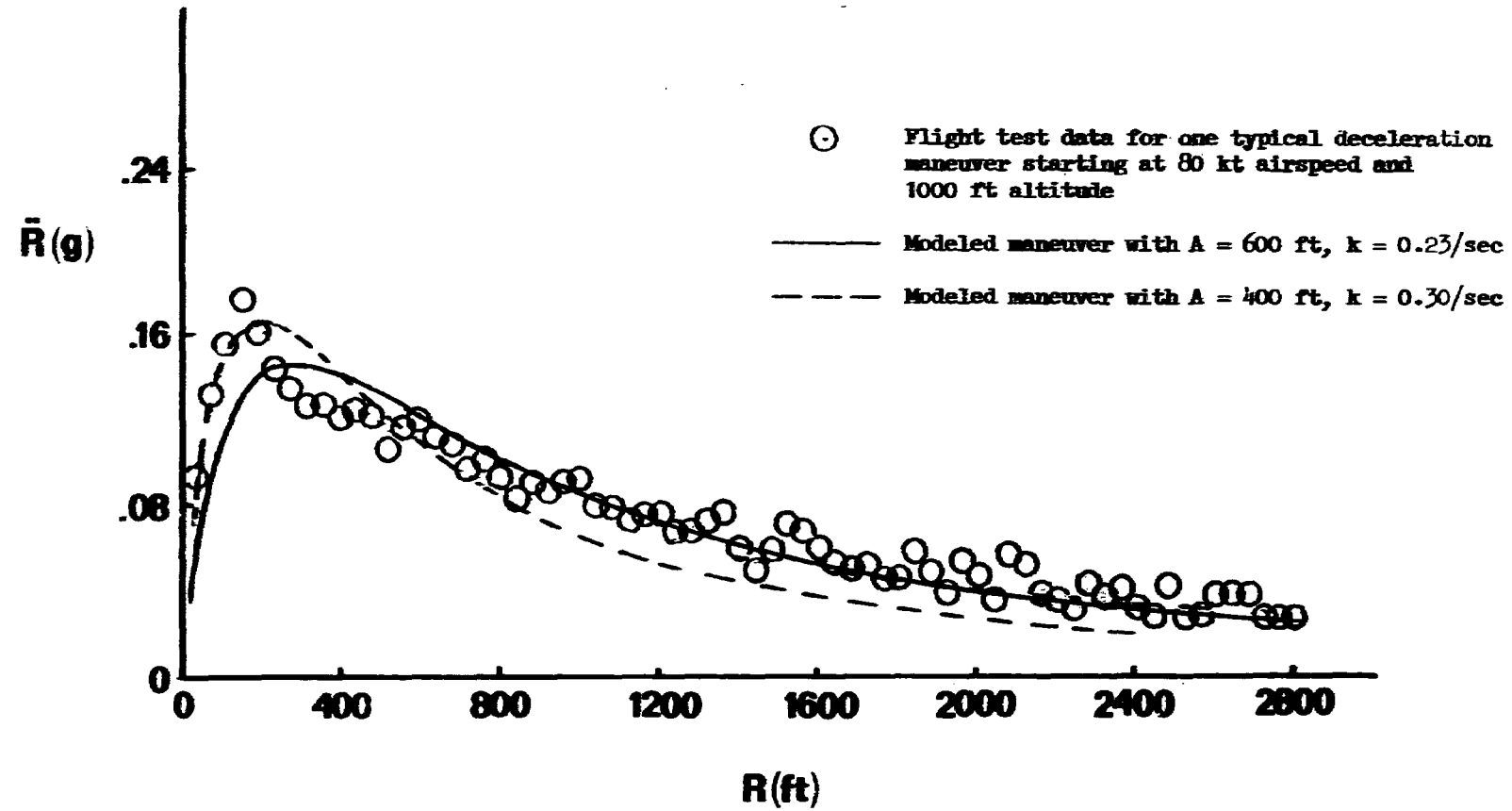


Figure 2. Comparison of Deceleration Profiles Between Analytical Model and Flight Test Data

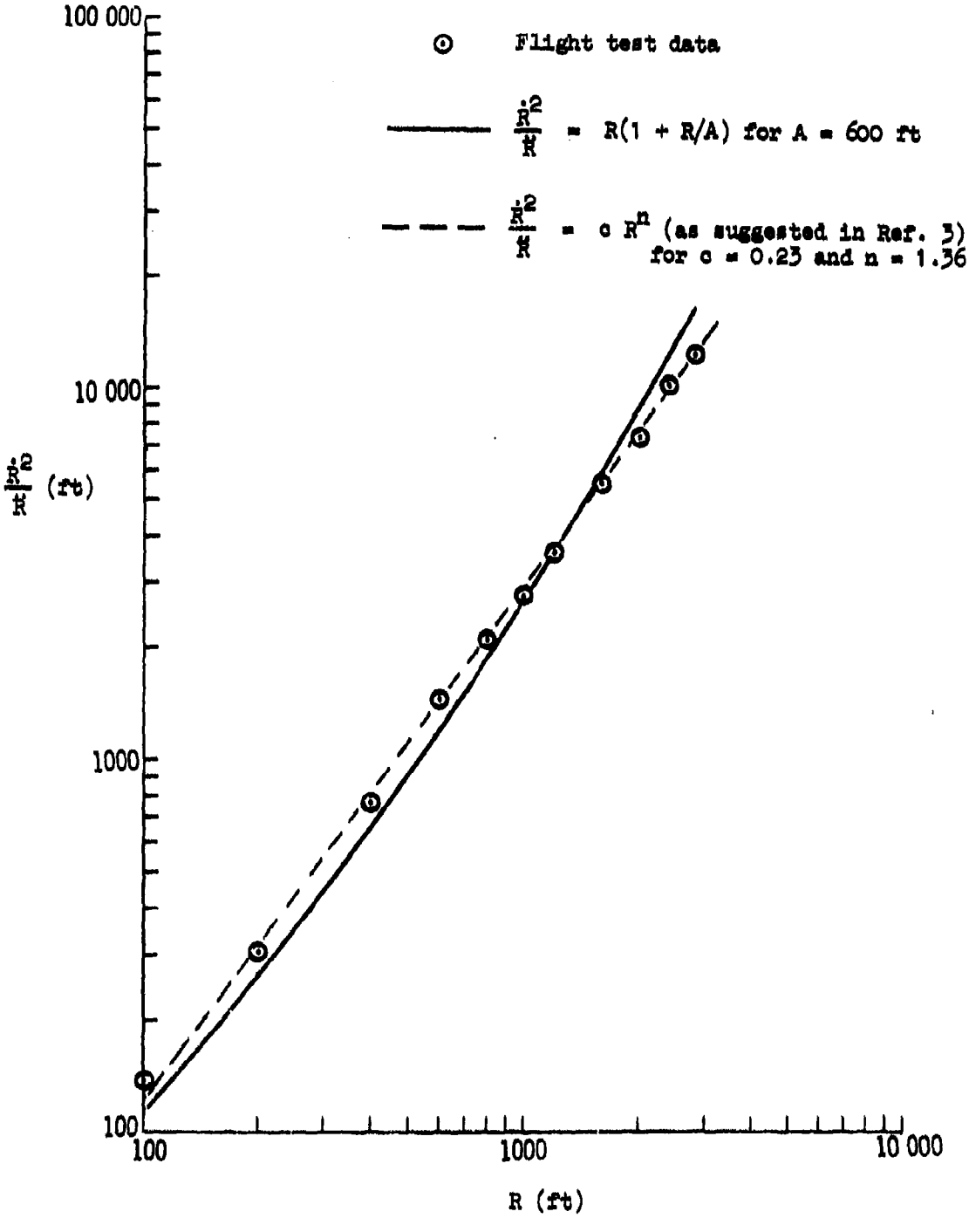


Figure 3. Logarithmic Plot of Parameter $\dot{R}^2/\ddot{R}$ Against Range

USEFULNESS OF THE MODEL

The deceleration model described here is useful because it represents pilot behavior over a wide range of speeds with a minimal number of parameters. At long range, the model yields a constant closure speed; at zero range it provides suitable hover position regulation.

The value of this is that it allows us to study a relatively complex and fairly long term maneuver with some insight into the factors involved. Interestingly enough, the two parameters involved can each be associated with the two aspects of the maneuver — the perception of the visual field and the response of the overall pilot-vehicle combination. For example, it may be desirable to relate the size of A to a specific landing site, e.g., to a conventional helicopter pad, to a shipboard landing pad, or to an open field. As for the pilot vehicle gain, k, we can readily identify its role as the crossover frequency in hovering. The model as formulated, however, implies that the effective crossover frequency is really range-varying according to how the landing site is visually perceived.

As a simulator tool the model described here has a special value in manned simulation. Decelerating approaches made on a simulator could be compared to actual approaches in terms of the two parameters identified. For example, according to Ref. 4 there is some evidence that the pilot's perceived range differs depending upon the means of displaying outside visual information in a simulator (i.e., whether a video display is collimated or not). The idea then would be to use this model as a simulator tool — to see whether the k and A of a simulation agree with those of a flight situation.

The model has a similar value with regard to training or establishing progress along a learning curve. It has been observed that as piloting skill develops for a given task, behavior becomes more consistent and starts to fit into a rather constrained pattern or model. This phenomenon has been observed in such tasks as glide slope tracking and landing flare.

One important application of the model may be with regard to automatic approach guidance systems. Since the model reflects the essential features of a manual decelerating approach, it could provide the basis of an autopilot guidance law which closely mimics manual operation. Thus the resulting automatically performed maneuver might be easier to monitor and more amenable to manual takeover, if required, than for guidance laws based on, say, constant deceleration or on velocity varying in direct proportion to range (exponential deceleration).

Finally, we speculate that the deceleration maneuver performed in an aircraft is akin to a stopping maneuver in an automobile, hence the analytical model presented here should apply. In stopping a car, most of us can probably identify with a gradual application of brakes (i.e., deceleration) up to a point then letting off the brake in the final few feet. Presumably, the effective size of the object being approached in stopping an automobile (i.e., the value of A) would be in line with the dimensions of a roadway or intersection.

EXTENSION OF THE MODEL TO OTHER AXES

The deceleration model as described here applies to range only. We can extend it to vertical and lateral axes, however, by the same application of rules of visual perspective.

According to Ref. 2 perceived size obeys the same relationship as perceived range in Eq. 2. We can follow this line of reasoning and hypothesize that:

$$\frac{\text{Perceived Altitude}}{\text{Actual Altitude}} = \frac{\text{Perceived Lateral Offset}}{\text{Actual Lateral Offset}} = \frac{1}{1 + R/A}$$

Hence,

$$\dot{h} = -k_h h_p = -k_h \frac{h}{1 + R/A}$$

and

$$\dot{y} = -k_y y_p = -k_y \frac{y}{1 + R/A}$$

Closed form solutions for velocities and accelerations can be obtained as functions of range in the same way as for range-related variables $\dot{R}$ and $\ddot{R}$. Although the detailed development and correlation with flight data is the subject of a subsequent paper, the altitude guidance model does appear to reflect the essential characteristics of actual decelerating approaches in helicopters. The usefulness of this extension follows that of the range model.

CONCLUSIONS

The fixed parameter deceleration model presented here is based on a rational hypothesis which combines the basic rules of visual perception with the idea of a constant coefficient crossover model of the human operator. The resulting guidance model formulation permits easy manipulation and solution using range as the independent variable. Most important, the model reflects the essential characteristics seen in actual flight maneuvers.

The usefulness of the visual deceleration model applies to a number of applications. In simulation, the model represents a validation tool as well as a gauge of pilot performance and learning. Since only two parameters are involved, on-line identification can be carried out with minimal computational impact. Also potential flight test applications are similar to those of the simulator. Even an automotive application may exist. But perhaps the most fruitful use of the model is in connection with automatic guidance law formulation where the mimicking of manual operation offers advantages for pilot monitoring. Finally, the same ideas used to create the range model can be applied to vertical and lateral guidance.

REFERENCES

1. McRuer, Duane T., and Ezra S. Krendel, Mathematical Models of Human Pilot Behavior, AGARD-AG-188, January 1974.

2. Gilinsky, Alberta S., "Perceived Size and Distance in Visual Space," Psychological Review, Vol. 58, 1951, pp. 460-482.

3. Moen, Gene C., Daniel J. DiCarlo, and Kenneth R. Yenni, A Parametric Analysis of Visual Approaches for Helicopters, NASA TN D-8275, December 1976.

4. Palmer, Everett, and John Petitt, "A Measure of Psychological Realism on a Visual Simulator," Journal of Aircraft, Vol. 14, No. 5, May 1977, pp. 421-422.

PREDICTING FIELD-OF-VIEW REQUIREMENTS FOR V/STOL AIRCRAFT APPROACH AND LANDING*

Warren F. Clement
Principal Research Engineer
Systems Technology, Inc.
2672 Bayshore-Frontage Road
Mountain View, California 94043

ABSTRACT

A method for quantifying fixed wing V/STOL aircraft field-of-view requirements is described and illustrated by a variety of examples involving different approach trajectories leading to a station-keeping point for shipboard recovery The analytic technique employed accounts for the combined variability in field of view caused by several factors. The first of these is the combination of six degrees of freedom of the aircraft as the pilot attempts to guide and to control the aircraft over the recovery area. Further, this technique includes the effects of the moving ship, airwake turbulence, limited visibility, and divided attention among multiloop control activities. The required field of view is a function of (1) the location of the pilot's point of regard on the ship, (2) the aircraft orientation, and (3) the relative position of the aircraft with respect to the ship, but in practice the field of view may be restricted by cockpit occlusions. The resulting field-of-view requirements depend in a critical way on the predicted model of the ship's airwake disturbance environment. Validation of the airwake environment for V/STOL operations with aviation facility ships, therefore, is an imperative recommendation resulting from this study. If airwake turbulence should prove to be as upsetting in reality as the present model suggests, the aircraft will need high authority, high bandwidth flight control systems in order to function effectively in this environment — much higher than is typical of current helicopter practice.

SYMBOLS

h	Height above station-keeping point
LOS_V	Depressed elevation angle of line of sight to center of landing pad with respect to horizon, except in Fig. 8, where LOS_V is defined to the upper edge of the hangar face. (This scale is not to be used in determining the subtended angle of any other point on the pad.) LOS_V is positive down.

* This research was sponsored by the Crew Systems Division of the Naval Air Development Center under Contract N62269-77-C-0509. Mr. Paul Linton was the contract technical monitor.

R_m	Observer's slant range to station-keeping point 40 ft above center of landing pad (70 ft above sea level)
t	Time to go to station-keeping point
WOD	Wind-over-deck
x_s	Ship's centerline axis positive forward of station-keeping point
y_s	Ship's lateral axis positive to starboard
$\Delta\psi$	Relative bearing angle of line of sight to center of landing pad with respect to aircraft nose, positive right. (This scale is not to be used in determining the subtended angle of any other point on the pad.)
λ, μ	Horizontal and vertical viewing angles, respectively, with reference to the nominal line of sight defined by $\Delta\psi$, LOS_γ. (These scales may be used to determine the subtended horizontal and vertical angle of any point on the pad). λ is positive right; μ, positive down.
σ	Root-mean-square deviation in general
$\sigma_\lambda, \sigma_\mu$	Horizontal and vertical semi-axes, respectively, of root-sum-square ellipses describing field of view perturbations due to aircraft and ship perturbations

INTRODUCTION

The most difficult piloting tasks encountered in vertical/short take-off and landing (V/STOL) aircraft are the transition from aerodynamically supported flight to thrust-supported flight, and the subsequent recovery aboard a ship or unprepared forward site. A significant portion of the difficulty experienced in these piloting tasks is due to the fact that out-of-the-cockpit visibility in past and current V/STOL designs has generally been too restricted to provide adequate visual cues. Aircraft designers have had difficulty in treating field-of-view problems since neither clear criteria nor standard procedures exist for determining required outside visibility for this class of aircraft.

OBJECTIVE

This paper develops a rationale for quantitatively determining fixed-wing V/STOL field-of-view requirements. It provides a relevant predictive and evaluative tool which models the complex interaction among human visual processes, the vehicle approach profile, and the pilot's flight path control performance. The model is intended to complement and support avionic system and crewstation design models. Applied to the development of a new air weapons system this

technique can significantly affect avionic and airframe design. The model specifies precise visual requirements for recovery aboard defined shipboard pads or forward sites. Those visual requirements which cannot be satisfied by piloting or canopy/window considerations must be fulfilled by innovative control and display concepts.

TECHNIQUE

The analytic technique employed involves the dynamic modeling of (a) the recovery guidance and control situation, (b) the disturbance environment, (c) the augmented aircraft, (d) the pilot's multiloop control activities, (e) the perceptual behavior of the pilot, and (f) the resulting geometric properties of information elements within the visual field. These factors are combined to identify those areas of the field of view which are essential for perception of the cues necessary for guiding and controlling the powered-lift aircraft to a safe recovery. The technique also accounts for the combined variability in the field of view caused by rolling, pitching, yawing, surging, swaying, and heaving degrees of freedom of the aircraft as the pilot attempts to guide and to control the aircraft over the recovery spot on a fixed or moving landing area. The required field of view is a function of (a) the location of the pilot's point of regard on the landing area, (b) the aircraft orientation, and (c) the relative position of the aircraft with respect to the desired landing point. In practice, the field of view may be restricted by cockpit occlusions; the technique will also identify and display these restrictions for a specified aircraft design.

The appendix describes the methodology and the references which detail its application to several nominal recovery approach trajectories to a destroyer of the USS Spruance class (DD963 through 966, according to Refs. 1 and 2).

CLASSIFICATION OF NOMINAL APPROACH TRAJECTORIES

Five special kinds of decelerating trajectories are considered here:*

- A homing or pursuit trajectory in which the aircraft flight path vector is always pointed toward the ship
- A trajectory which follows a straight line path in the earth reference frame toward a predetermined intercept point with the ship

---

* A sixth variable deceleration trajectory (at constant glide slope and horizontal bearing) in which inertial velocities tangent and normal to flight path are proportional to visually perceived position — a trajectory which closely approximates manual, head-up visual approaches to hover — is described in a companion paper, Ref. 3.

- A trajectory which maintains a constant bearing with respect to the ship (such as a conventional approach using fixed line-up and glide slope angles)

  - A constant sink rate trajectory (at constant horizontal bearing) where deceleration along the earth x-axis is constant

  - A constant altitude trajectory (at constant horizontal bearing) where deceleration along the earth x-axis is constant

Figure 1 depicts examples of the five vertical plane trajectories and three horizontal plane trajectories in the ship's frame of reference. The ship is underway at 20 kt on a fixed course in calm air. Thus the forward motion of the ship is generating the relative wind-over-deck (WOD) of 20 kt along the ship's centerline. (Subsequently we shall estimate the disturbing effects of the ship's motion and air wake turbulence in sea state 5 with 43 kt WOD.) Each trajectory terminates at 40 ft deck height directly over the center of the recovery circle. The termination of the homing or pursuit trajectory is aligned with the WOD and coincides with the constant altitude trajectory by design without special maneuvers, but the missed approach route is blocked by the aft mast and stack. Both the straight line (in the earth reference frame) and the constant sink

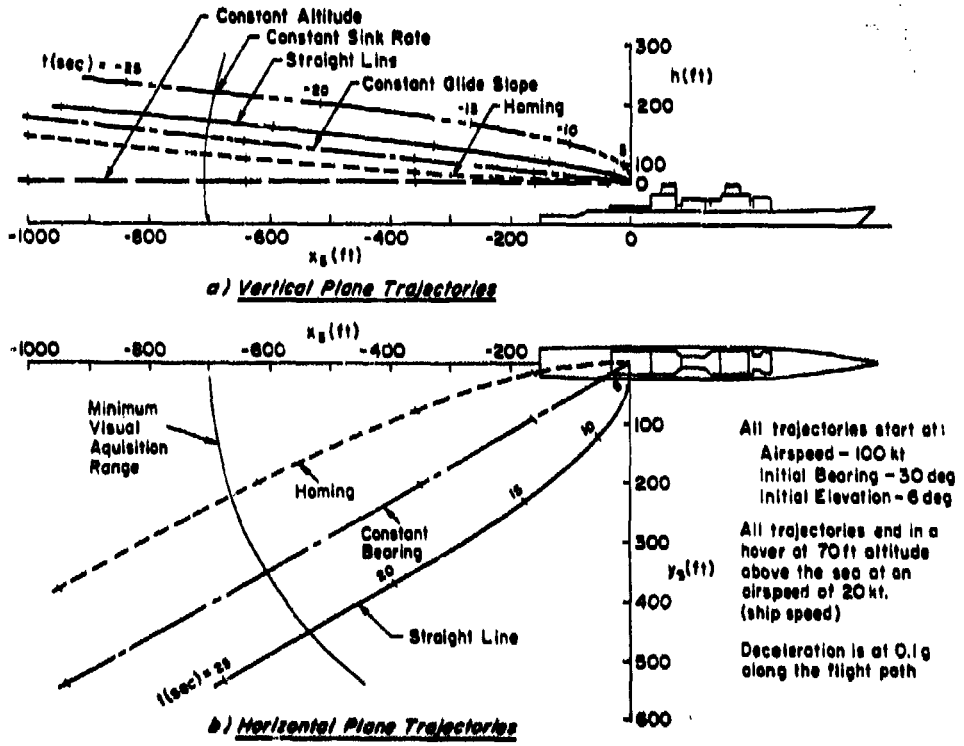


Figure 1. Exemplary Approach Trajectories in Ship's Reference Frame.

rate trajectories require special terminal maneuvers to arrest the descent and to match the ship's velocity. The constant bearing and glide slope trajectory (in the ship's reference frame) is compatible with existing visual landing aids on the ship. Maneuvering requirements for the homing or pursuit and constant bearing trajectories are discussed in Ref. 4.

THE PILOT'S VIEW FROM EACH NOMINAL TRAJECTORY WITHOUT AIRCRAFT PERTURBATIONS

Nominal field of view requirements will be shown for a V/STOL aircraft on final approach to the recovery deck of a Spruance-class destroyer by a sequence of perspective views of the ship from the aircraft beginning at the minimum visual acquisition range, 700 ft, identified in Fig. 1. Each sequence ends with the pilot's eye 40 ft above the center of the recovery circle.

Three types of approach trajectories will be considered first: homing or pursuit, straight line (in the earth frame), and constant bearing (in the ship's frame). In addition, a fourth constant sink rate trajectory and a fifth "constant altitude" trajectory in the vertical plane will be combined with a constant bearing trajectory in the horizontal plane. Detailed descriptions of these trajectories can be found in Ref. 5.

The results of the undisturbed field of view calculations for the nominal homing, straight line, constant bearing, and constant sink rate trajectories are presented in Figs. 2 through 6, respectively. The illustrations assume that the pilot is looking at the center of the recovery circle, which appears elliptical in perspective; the figures omit, for the time being, the obscuring effect of any cockpit structure. The origin of the horizontal viewing angle (λ) axis corresponds to the horizontal reference line of sight relative to the nose of the aircraft at each value of slant range, R_m; thus the centers of the landing pads appear to be in the same horizontal line of sight for the homing and constant bearing trajectories, but not for the straight line trajectory. The origin of the vertical viewing angle (μ) axis corresponds to the vertical line of sight; thus the centers of the landing pads reflect the depressed elevation angles (LOS_V) relative to the horizon at each slant range, R_m.

As long as the aircraft (pitch and roll) attitude is approximately level, field-of-view requirements for the nominal unperturbed approach trajectories may be interpreted directly in terms of the angles λ and $\mu+LOS_V$ for all trajectories except the straight line (in inertial space), for which the horizontal field-of-view requirement must be interpreted in terms of $\lambda + \Delta\psi$.

Figures 7 and 8 are projections of the landing pad and hangar face, respectively, when the aircraft has reached the hover point 40 ft over the center of the pad ($R_m \doteq 0$). In Fig. 7 the pilot is assumed to be looking at the center of the recovery circle which is, of course, at the nadir, but in Fig. 8, his line-of-sight (LOS) is shifted to the upper center of the hangar face. Even that line of sight is depressed over 33 deg from the horizontal plane.

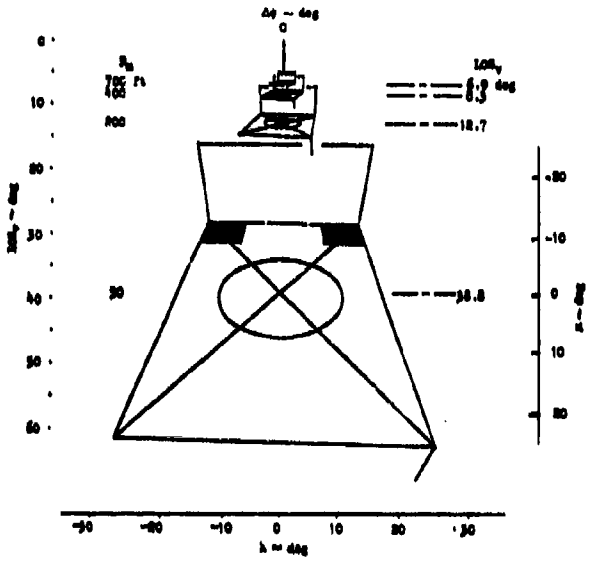


Figure 2. Perspective Views of Landing Pad for a Homing Trajectory

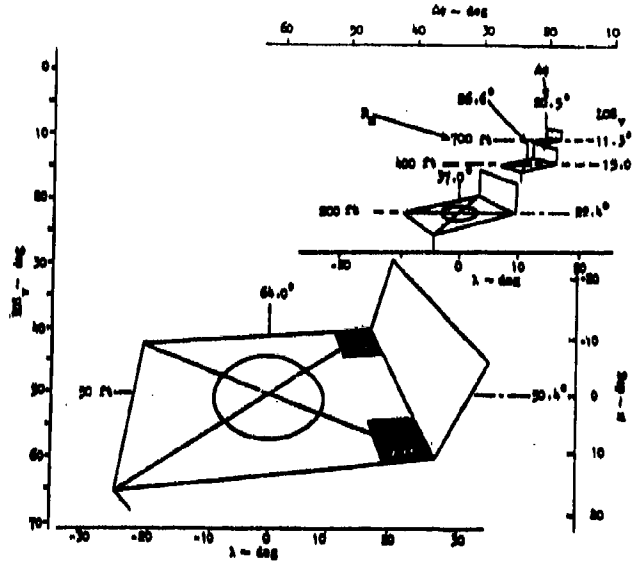


Figure 3. Perspective Views of Landing Pad for a Straight Line Trajectory

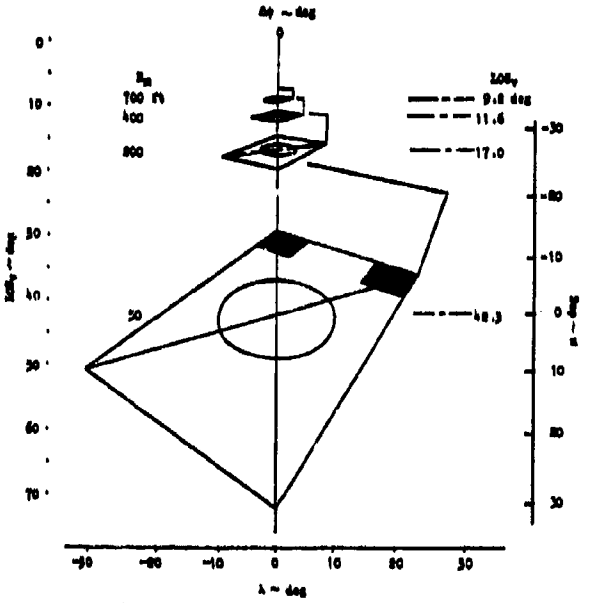


Figure 4. Perspective Views of Landing Pad for a Constant Bearing Trajectory

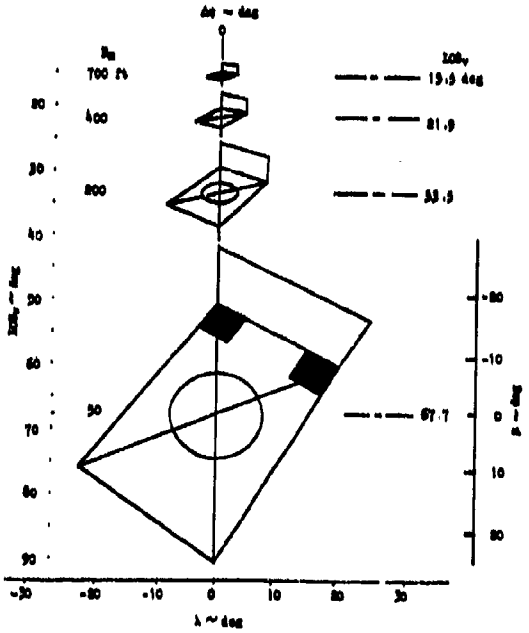


Figure 5. Perspective Views of Landing Pad for a Constant Sink Rate Trajectory at Constant Bearing

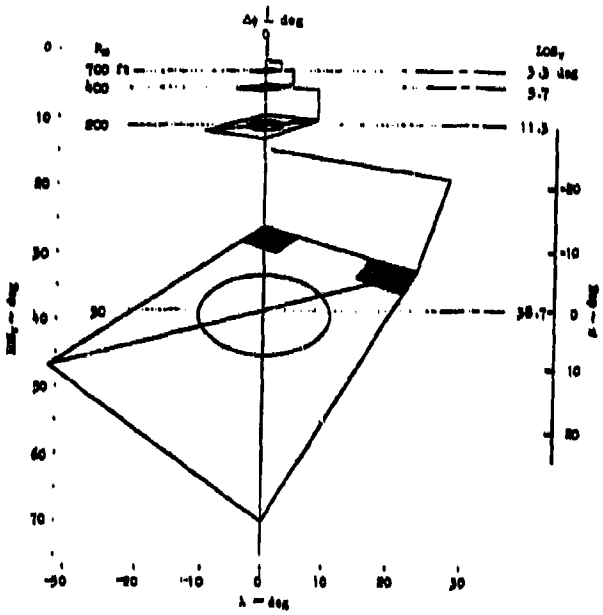


Figure 6. Perspective Views of Landing Pad for a Constant Altitude Trajectory at Constant Bearing

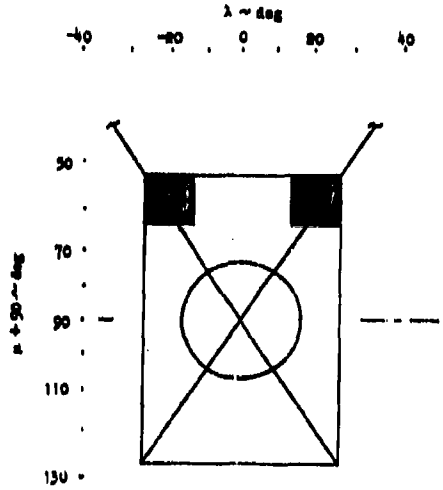


Figure 7. Perspective View of Landing Pad for an Observer 40 ft Above Landing Pad and Line of Sight Through Center of Pad

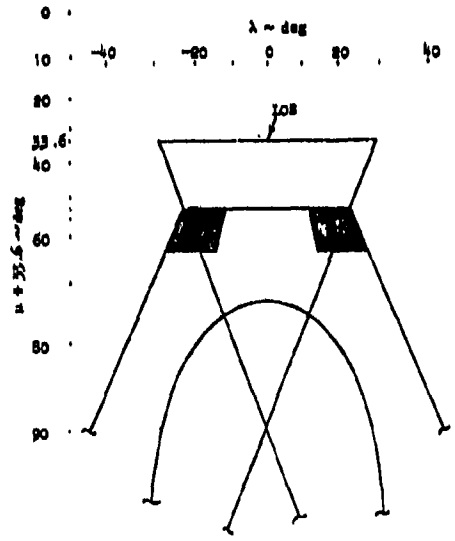


Figure 8. Perspective View of Landing Pad for an Observer 40 ft Above Landing Pad and Line of Sight Through Upper Center of Hangar Face

Table 1 summarizes some tentative conclusions based on the foregoing examples of the pilot's view from each nominal trajectory without aircraft perturbations.

TABLE 1

COMPARATIVE CONCLUSIONS ABOUT FIELD OF VIEW AMONG SEVERAL NOMINAL APPROACH TRAJECTORIES WITHOUT AIRCRAFT PERTURBATIONS

Homing trajectory (Fig. 2)

Vertical field is smallest since angle over the nose is least until over the deck.

Lateral field is smallest, if sideslip angle is zero.

Straight line trajectory (Fig. 3)

Vertical field is larger, but ship may not disappear from view as early as in constant sink rate.

Lateral field is greatest, if sideslip angle is zero.

Constant bearing trajectory (Fig. 4)

Vertical field is intermediate, between that for homing and straight line until reaching 50 ft range-to-go

Lateral field is smallest, if aircraft heading is maintained constant.

Constant sink rate trajectory (Fig. 5)

Vertical field is greatest — ship may disappear from view below nose at considerable range.

Lateral field is same as for constant bearing.

Constant altitude trajectory (Fig. 6)

Vertical field is approximately the same as for homing.

Lateral field is same as for constant bearing.

In the next topic we shall illustrate the additional field-of-view requirements imposed by controlled aircraft perturbations caused by the relatively severe disturbance environment associated with sea state 5 and 43 kt wind-over-deck.

EXAMPLES OF THE PILOT'S VIEW FROM 50 FT RANGE TO THE HOVERING POINT WITH CONTROLLED AIRCRAFT PERTURBATIONS

In the previous topic we have illustrated how the location of the point of regard on the ship and how the relative position of the undisturbed aircraft with respect to the ship affect the pilot's field of view. We have seen that the field-of-view requirements will be greatest near the ship in hovering and near-hovering flight. We now illustrate the combined effects of rotational and translational disturbances in the controlled aircraft motion on the hovering and near-hovering field-of-view requirements.

The details of the extensive analyses required to provide the examples shown in this topic are documented in Refs. 6 and 7 and will not be repeated here.* The results depend in a critical way on the predicted model of the ship's airwake disturbance environment adopted in Ref. 6 for sea state 5 and 43 kt wind-over-deck, which present quite a severe disturbance environment. Two examples of perturbed field-of-view requirements are presented: (1) from 50 ft range-to-go (to the hovering point) on a constant relative bearing trajectory and (2) from the hovering point itself 40 ft above the center of the deck. For selected points of regard by the pilot we can estimate Gaussian statistical variances in the required field of view caused by the predicted variances in the six aircraft degrees of freedom listed in Table 2.† Corresponding distributions of perturbations in terms of viewing angles will, however, tend to be increasingly skewed as the point of regard departs from the nominal line of sight (the optical axis of the picture plane), because of the tangent relationship between picture plane coordinates and viewing angles. Nevertheless we can approximate the root-sum-squared angular variance for the pilot's nominal line of sight to the center of the recovery circle by the large ellipse in Fig. 10 at 50 ft range-to-go to station-keeping. The root-sum-squared variance

---

* A method for introducing cockpit occlusions into the field-of-view analysis is also presented in Ref. 7. An example of the cockpit occlusion for the AV-8A Harrier is shown in Fig. 9 with the nominal field of view of the ship superimposed for 50 ft range-to-go to station-keeping from Fig. 4. Clearly this example indicates field of view problems even without the effects of disturbed vehicle motions.

† The predicted aircraft motion variances in Table 2 are based on closed-loop pilot-vehicle analysis of the disturbed aircraft in station-keeping flight tracking the ship's landing pad, which is, in turn, disturbed by the sea. Predicted ship's motion variances at the center of the landing pad are also listed in Table 2. The pilot-vehicle analysis in Ref. 6 used an intermittent or divided attention model for the pilot's remnant together with the following crossover frequencies typical of unaided manual aircraft control.

Control Loop	Crossover Frequency (rad/sec)
roll	1.5
pitch, yaw	1.0
heave	0.3
surge, sway	0.2

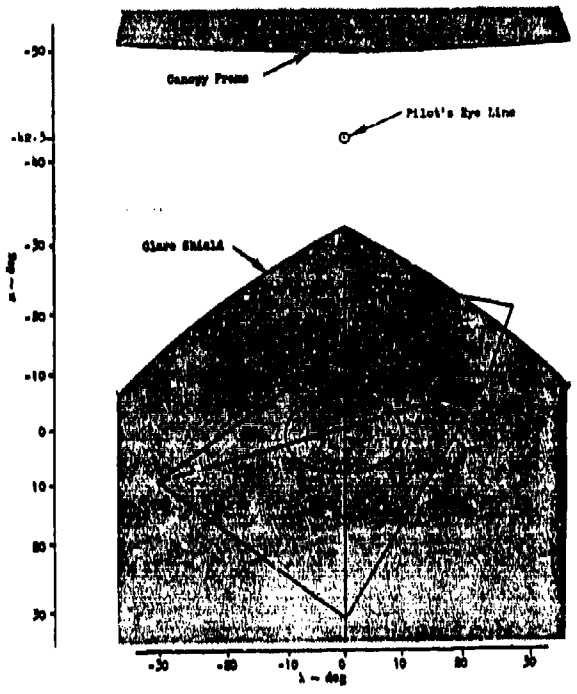


Figure 9. Binocular Projections of the AV-8A Cockpit Occlusions Onto the Viewing Plane for R_m = 50 ft

TABLE 2

SUMMARY OF ILLUSTRATIVE MOTION ERROR VARIANCES
FOR USE IN EXAMINING HOVERING FIELD-OF-VIEW REQUIREMENTS
(From Ref. 6)

Sea State 5

Wind-Over-Deck = 45 kt

Ship Speed, V_s = 20 kt

Clockwise Wave Direction Relative to Ship's Stern, μ = 120 deg

Mean-Squared Ship Motions at the Center of the Landing Pad (Relative to Inertial Space) for the DL965

Ship Motion		Variance
Roll	σ_φ^2	$(2.02 \text{ deg})^2$
Pitch	σ_θ^2	$(0.77 \text{ deg})^2$
Yaw	σ_ψ^2	$(0.30 \text{ deg})^2$
Surge	σ_x^2	$(0.40 \text{ ft})^2$
Sway	σ_y^2	$(2.05 \text{ ft})^2$
Heave	σ_z^2	$(2.71 \text{ ft})^2$

Mean-Square Aircraft Motion Relative to Ship while Stationkeeping under Manual Control

Aircraft Motion		Variance
Roll	σ_φ^2	$(2.5 \text{ deg})^2$
Pitch	σ_θ^2	$(3.0 \text{ deg})^2$
Yaw	σ_ψ^2	$(2.1 \text{ deg})^2$
Surge	σ_x^2	$(11.6 \text{ ft})^2$
Sway	σ_y^2	$(15.5 \text{ ft})^2$
Heave	σ_z^2	$(19.5 \text{ ft})^2$

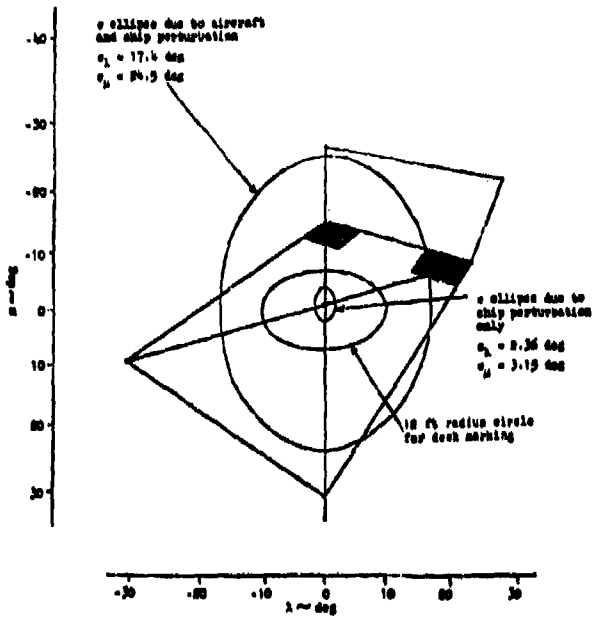


Figure 10. Root-Sum-Square Variation for the Center of the Pad in the Pilot's Field of View for R_m = 50 ft

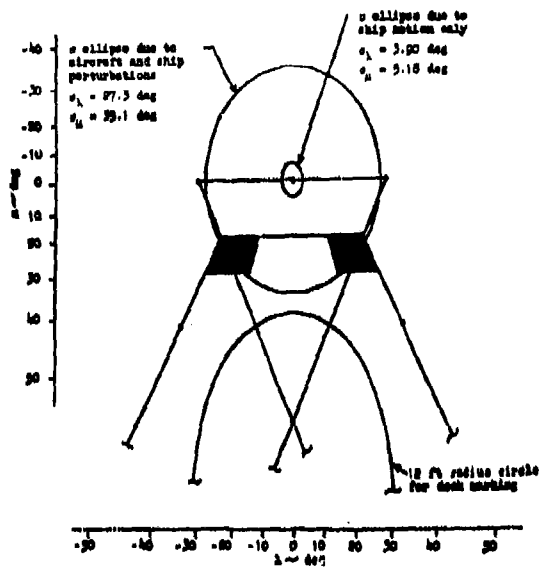


Figure 11. Root-Sum-Square Variation for the Center of the Upper Edge of the Hangar in the Pilot's Field of View for R_m = 0

for the pilot's nominal line of sight to the upper center of the hangar face is shown by the large ellipse in Fig. 11 while the pilot is attempting to keep on station over the deck. These extreme estimates of variability confirm the desirability, from the standpoint of reducing field-of-view requirements, of providing visual cues and visual aids above the hangar for assisting the pilot in arresting the approach and providing hovering guidance.

These extreme estimates of variability also emphasize the need for high authority, high bandwidth flight control systems in order to function effectively in this environment — much higher than is typical of current helicopter practice. The allowable aircraft motions due to all causes are only a few feet — Ref. 8 states "3 to 4 ft" in "allowable" touchdown error. If we interpret "allowable" as 2σ, Table 2 shows that the square root of the sum of the x- and y-variances in deck motion alone at the center of the pad slightly exceeds this "allowable" touchdown error. Without chasing the deck, this allowable touchdown error thus represents the best precision one might expect from a guidance and control system which regulates so well that no errors are contributed by aerodynamic disturbances or the pilot's divided attention. Graphic representation from the pilot's station-keeping perspective of this apparent σ-variation in the center of the pad due to ship motion alone is shown by the smallest central σ-ellipse at 50 ft range-to-go in Fig. 10 and at zero range-to-go in Fig. 11.

With this assessment of the variability in the viewing angles required of the pilot with respect to his cockpit reference, we conclude our analyses of field-of-view requirements in the extremely severe disturbance environment to be expected in sea state 5 with 43 kt wind-over-deck. The predicted field-of-view requirements with manual control appear to be so great that more confidence is needed in the validity of the predicted airwake disturbance environment on which this analysis is predicated before these expectations can be converted into design requirements for field of view which are, in turn, based on high authority, high bandwidth flight control systems.

SUMMARY

The field of view (FOV) from the aircraft cockpit required to see a point of regard on the recovery area during an approach to hovering over the area is a function of (a) the location of the point of regard itself on the ship, (b) the aircraft orientation, and (c) the relative position of the aircraft with respect to the ship, but the FOV may be restricted by cockpit occlusions.

We have illustrated how the location of the point of regard on the ship and how the relative position of the undisturbed aircraft with respect to the ship affect the pilot's field of view from several types of approaching trajectories. The field-of-view requirements will be greatest near the ship in hovering and near-hovering flight above the recovery area.

In addition to the undisturbed aircraft field-of-view requirements, we have illustrated how the effects of rotational and translational disturbances in the controlled aircraft motion combine to increase further the hovering and near-hovering field-of-view requirements.

Two examples of perturbed field-of-view requirements are presented: (1) from 50 ft range-to-go (to the hovering point) on a constant relative bearing trajectory and (2) from the hovering point itself 40 ft above the center of the landing pad. In hovering at zero range-to-go, the vertical field of view required to observe (plus or minus) one standard deviation of the upper port corner of the hangar would be -34 ± 20 deg over the nose, assuming maximum sideslip and horizontal pitch attitude. The corresponding (plus or minus) standard deviation in the lateral field of view would be ± 19 deg. If instead, one were to assume no sideslip, the same vertical field requirement would apply, but the lateral requirement would become -28 ± 19 deg (negative left of the nose for an approach from the starboard quarter of the ship). In addition, while hovering at zero range-to-go, the vertical field of view required to observe (plus or minus) one standard deviation of the most forward line-up light would increase to -53 ± 25 deg, but the lateral requirement would become -22 ± 22 deg, assuming no sideslip.

Finally, even with 50 ft of range still to go on the constant bearing trajectory, the vertical field of view required to observe (plus or minus) one standard deviation of the center of the landing pad would be -43 ± 24 deg, but the lateral requirement would become -20 ± 17 deg, assuming no sideslip.

These extreme estimates of variability confirm the desirability, from the standpoint of reducing field-of-view requirements, of providing visual cues and visual aids above the hangar for assisting the pilot in arresting the approach and providing hovering guidance.

CONCLUSION

If airwake turbulence should prove to be as upsetting in reality as the present model suggests, the aircraft with its control system needs to achieve higher crossover frequencies in translational degrees of freedom — higher than is typical of current helicopter practice. Depending on technique, this may also require higher crossover frequencies in roll and pitch axes. This may require a careful tradeoff of the roles of manual and automatic control in order to keep the pilot's intermittent/divided attention noise from compromising precision in station-keeping.

APPENDIX

PROCEDURE FOR DETERMINING EXTERNAL VISUAL FIELD-OF-VIEW REQUIREMENTS

This appendix outlines the procedure for determining external visual field-of-view requirements. Each of the four major steps in the procedure will be described briefly and partitioned into subsidiary tasks, for which there are six supporting working papers to help the interested analyst in carrying out the technical details. The six working papers are identified by reference number. The interactive procedure for determining field-of-view requirements presently also includes two small HP-67/97 computer programs supplied with the corresponding working paper to aid the analyst. The six working papers can be obtained at cost from the technical librarian of Systems Technology, Inc.

Step I — Kinematic Representation of Aircraft Approach Trajectories to a Landing Pad on a Moving Ship or Shore-Based Site (Ref. 9)

The first step in the procedure is to define the intended approach trajectory of the aircraft with respect to the recovery area on a moving ship or shore-based site. The pertinent coordinate transformations and kinematic relationships for a point-mass aircraft are employed. The results of this step include the sequence of nominal operating points in terms of range, bearing, and altitude with respect to the terminal station-keeping or hovering point over the pad as a function of time. The specific tasks for this step require that the analyst:

a. Specify ship velocity, if appropriate; specify desired type of approach trajectory, deceleration, and terminal station-keeping or hovering coordinates of aircraft.

b. Define position coordinates of a point-mass aircraft relative to the terminal hovering point over the pad as a function of time (HP-67/97 program is available for constant relative bearing, rectilinear, and constant sink rate; Tymshare PDP-10 program, for homing).

Step II — Pilot/Vehicle Performance

The second step in the procedure is to estimate what effects pilot-controlled vehicle (rigid-body) motions about the vehicle center of mass will have on perturbations in following the sequence of operating points during the course of a typical approach. The procedure has been described in Refs. 10 and 11. This requires the quantification (Ref. 12) and exercising (Ref. 6) of an analytical model of the pilot-plus-vehicle executing the intended approach trajectory under the influence of turbulence, ship motion, degraded visibility, and pilot variability (e.g., as caused by divided attention to several control and monitoring tasks during the approach). The results of this task are quantitative relationships between the trajectory variables, ship motions, turbulence, and piloting noise on the one hand and the variability in the angular and translational kinematics of the aircraft on the other. The results of Step II are described in Ref. 6.

Step III — Visual Element Motion Analysis

The third step in the procedure is to convert the sequence of operating points along the approach trajectories, as well as the perturbations about these trajectories, into corresponding apparent angular kinematics ascribed to a set of essential visual elements in the pilot's external field of view. In this study the essential visual elements are the deck pad and hangar delineation for a DD963 class of destroyer (Refs. 1 and 2), which has a 40 ft-by-60 ft recovery area and a 40 ft wide-by-20 ft high hangar face. The results of Step III are presented in Ref. 5. The specific tasks for Step III require that the analyst:

a. From the sequence of nominal trajectory operating points in Step I, identify geometric forms (and their sensitivity functions) of essential elements or referents in the visual field which provide cues for outer loop task variables as a function of meteorological conditions affecting contrast and visual acuity (Ref. 13).* Reference 1 may be consulted for details of recovery aids and area geometry.

b. By means of the necessary coordinate transformations and kinematic representations, relate nominal aircraft trajectory motions and positions in Step I to angular motions and positions of essential referents in the pilot's forward visual hemisphere (Ref. 14 and HP-67/97 program).

---

* Observation thresholds are based on contrast, visual acuity. Pilot indifference thresholds, in turn, are based on observation thresholds and tolerances on precision required for the task.

c. By means of perturbed linearized coordinate transformations and kinematic representations, relate disturbances in ship and aircraft motions and positions to variability in the nominal angular motions and positions of essential referents in the pilot's forward visual hemisphere (Ref. 7).

d. Superimpose angular measures of variability (e.g., root-mean-squares) in Step c on nominal motions and positions of essential referents in Step b (Ref. 7).

e. Partition contributions to variability from ship motion, atmospheric turbulence, and pilot-induced noise in Step II (Ref. 6).

Step IV — Field-of-View Requirements

Combining in Step IIId the "nominal" angular motions and positions of essential referents from Step IIIb with the estimates of variability in these same nominal motions and positions from Step IIIc allows the analyst to define explicitly the location of essential visual cues (for guidance and control of the aircraft) within the pilot's forward visual hemisphere from which field-of-view requirements can be established. Furthermore the source of each contribution, whether nominal or variable with some probability, can be identified. Thus, similarly appearing variations with different causes could be identified as potential sources of ambiguity in interpretation. The concluding assessments of field-of-view requirements from the examples examined in this fourth step are presented in this paper. The specific tasks in Step IV require that the analyst:

a. Establish probable bounds on the forward hemispheric field of view which includes the essential cues and referents from Step IIId.

b. Based on the apparent geometric sensitivity functions (Step IIIa), and the acceptable range of the pilot's gain adaptation, estimate the comparative usefulness of the available cues for the guidance and control of the aircraft; and identify essential areas within the probable bounds on the field of view from Step IVa for perception of guidance and control cues.

c. Estimate minimum field-of-view boundaries which include essential areas.

d. Compare results with typical cockpit occlusions and fields of view for head-up displays.

REFERENCES

1. Benedetto, R. A., Non-Aviation Ships Helicopter Facility Resume, NAEC ENG-7575, Rev. C., 1973.

2. Jane, Fred T., (Capt. John E. Moore, Ed.), Jane's Fighting Ships, McGraw Hill Book Co., N. Y., 1973-74, p. 420.

3. Heffley, Robert K., A Model for Manual Decelerating Approaches to Hover, Systems Technology, Inc., Paper 228R, Presented at the 15th Annual Conference on Manual Control, Wright State University, Dayton, Ohio, 20-22 March 1979.

4. Ringland, Robert F., Approach Trajectory Considerations for Terminal Operations of Navy V/STOL Aircraft, Systems Technology, Inc., Paper 227, Presented at the AIAA Guidance and Control Conference, Palo Alto, California, August 1978.

5. Clement, Warren F., Robert K. Heffley, and Wayne F. Jewell, Field-of-View Requirements for Approach and Landing of V/STOL Aircraft, NADC 77240-07, August 1978.

6. Heffley, Robert K., Calculation of Total Motion Error Variance for Hovering Field of View Analyses, Systems Technology, Inc., Working Paper 1115-5, July 1978.

7. Jewell, Wayne F., and Warren F. Clement, Perturbations of the Pilot's View of a Non-Aviation Ship's VTOL Pad Due to Linear and Angular Perturbations of the Aircraft, Systems Technology, Inc., Working Paper 1115-4, August 1978.

8. Miyashiro, S. K., and F. E. Morris, "VTOL/Helicopter Approach and Landing Guidance Sensors for Navy Ship Applications," Navy/NASA VSTOL Flying Qualities Workshop, August 1977, p. 497.

9. Heffley, Robert K., Computation and Analysis of V/STOL Approach Trajectories, Systems Technology, Inc., Working Paper 1122-1, March 1978.

10. Clement, W. F., D. T. McRuer, and R. H. Klein, "Systematic Manual Control Display Design," Guidance and Control Displays, AGARD CP-96, February 1972.

11. Clement, Warren F., R. Wade Allen, and Dunstan Graham, Pilot Experiments for a Theory of Integrated Display Format, JANAIR Report 711107, October 1971.

12. Heffley, Robert K., and Warren F. Clement, Development of General Closed Loop Pilot Vehicle Dynamics for VTOL Aircraft at Low Speeds, Systems Technology, Inc., Working Paper 1115-1, April 1978.

13. Clement, Warren F., and Robert K. Heffley, Some Effects of Adverse Visibility on Threshold Properties of the Pilot's Perception in VTOL Approaches to Non-Aviation Ships, Systems Technology, Inc., Working Paper 1115-3, May 1978.

14. Jewell, Wayne F., and Warren F. Clement, Computation of the Pilot's View of a Non-Aviation Ship's VTOL Pad for Several Types of Approach Trajectories, Systems Technology, Inc., Working Paper 1115-2, March 1978.

SESSION 5B: DISPLAY CONSIDERATIONS

Moderator: Dr. Malcolm L. Ritchie, Wright State University

5b

EXTENDED EVENT-DRIVEN DISPLAYS FOR MANIPULATOR CONTROL

by
Garrett Paine
Member of Technical Staff

and

Antal K. Bejczy
Member of Technical Staff

Jet Propulsion Laboratory
California Institute of Technology
Pasadena, California 91103

ABSTRACT

This paper deals with two extensions of event-driven display technology for proximity and force-torque sensor data: (a) Higher level automation for event data processing and display. (b) The operator's vocal interaction with the display controller. Both extensions are aimed at the performance enhancement of remote manipulator control systems. The paper describes the technical details of voice-commanding the display by a discrete word recognition system, and elaborates on the task-related algorithms driving the display. These algorithms include the capability of automatically changing the format or contents of the display to provide the operator with the most appropriate information on the task execution. The manipulator system in which the new display concepts are being tested includes both manual and supervisory control implemented on a dedicated mini- and microcomputer network. In computer-based supervisory control the displays help the operator evaluate how well the computerized control is being performed in the absence of specific, new supervisory inputs and to intervene, if necessary.

I. INTRODUCTION

The general problem of constructing event-related information displays from diverse and multi-dimensional data generated by proximity, force-torque and tactile sensors integrated with the end effector of a robotic arm has been treated in a previous paper (Ref. 1). Event-driven sensor information displays provide a convenient means by which the operator's attention can be focused on control goals or subgoals expressed in terms of sensor data. The implementation of event-driven displays requires the development and use of real-time algorithms which coordinate and evaluate sensor data based on predefined control events and drive an appropriate display conveying the event information to the operator.

Manipulator control tasks can be subdivided into subtasks which can be referenced to sensor data and defined as sensory events. Each event may have a variety of characteristic parameters. The development of general purpose event-driven displays requires, thus, that the display drive algorithms be flexible in the sense that changing control goals or subgoals can easily be accommodated by simple changes in the algorithmic parameters to match the needs of a given control task.

The development of event-driven displays requires the consideration of various technical and human factors issues. E.g., how to integrate different displays into a coherent working format? How much and what kind of detailed information the operator should be exposed to in addition to the event information? How should the operator control the displays?

Event-driven displays are expected to improve overall control performance for several reasons. First, they can simplify on-line control decisions. Second, they can reduce perceptive/cognitive workload on the human operator in a real-time control environment. Third, they can reduce errors caused by human factor effects.

Extensions to event-driven display techniques have been developed which provide more task-relevant information to the operator, reduce his workload, and give him better control of the displays. More task-relevent information has been provided to the operator by combining hand geometry data with the sensor data to show him both the sensor pictorial and the corrective control required to reach a desired end state. The operator's workload has been reduced through extension of the concept of event detection and annunciation to one where the displays' contents are automatically modified depending on task performance. Voice control of the display system state and parameters has been developed to permit the operator to modify the display without having to use his hands for display control or move his head to scan several displays for different information. These extensions thus enable teleoperated and sensor referenced manipulator tasks to be performed more reliably, more efficiently, and with a lower consumption of resources.

Section II describes the content and use of the new algorithms and display formats which extend event-driven display technology by employing higher level automation for sensor data processing. Section III describes the present status and use of voice controlled graphics displays. Conclusions are summarized in Section IV. Plans for future work are included in Sections II and III. Appendix A provides a brief description of the Advanced Teleoperator Development Laboratory at the Jet Propulsion Laboratory where this work was performed.

II. AUTOMATION TECHNIQUES FOR DISPLAY CONTROL

Successful manipulator control, in both manual and computer control modes, requires, in general, that the operator know: what is to be done, what the status of the task (man/machine/object/environment) is, and what control decisions must be made to complete the task. The terminal/compliance phase of

manipulator motion--when the mechanical hand is near to or in contact with the object--is characterized by a high density of diverse and rapidly changing sensor data which have to be evaluated in real time for successful control. Experimental studies indicate (Refs. 2 and 3) that 25-50% of typical manipulator control task time is spent in the terminal/compliance phase of manipulator motion and most errors and failures occur during this phase of motion control.

The information that the operator needs during the terminal/compliance phase of manipulator motion control includes arm geometry, arm geometry rates, object parameters, environmental parameters, hand configuration, contact forces and torques, contact area pressure distribution and slip once the object has been grasped, etc. Much of the operator's workload can be eliminated if he does not have to be responsible for sequencing displays, changing coordinate frames, or making fine quantitative interpretations from TV or graphics display data. Fortunately, appropriate event-driven displays can do many of these things for him automatically.

1. Encounter Regime Displays

Encounter regime display control problems come into effect when the mechanical hand is near to or in contact with the object to be handled or avoided. The encounter regime, a term chosen to emphasize the information aspects of the terminal/compliance phase of the manipulator motion, is shown diagrammatically in Figure 1. Here a manipulator is approaching a block resting on a table, and the operator must align the hand to the block prior to grasping it. Without the ability to randomly position TV camers, as distinct from pointing them, the operator will need precise information about the relative location of the hand to the table, and the relative location and orientation of the hand to the block. This quantitative data can be provided by hand based proximity sensors (Refs. 4 and 5). The dashed lines in Figure 1 indicate the lines of sight of four finger mounted proximity sensors.

The coordinates associated with the encounter regime are shown in Figure 2. The d_i's are the path lengths detected by the proximity sensors. As seen from the wrist of the manipulator the approach angle is equivalent to yaw and the hover angle to roll. By changing the separation between the fingers the d_1 and d_2 measurements can be used both to find the approach angle and find the sides of the object (block). Note that the singularity (s=0) is not reached in the latest JPL sensor configuration due to the location of the proximity sensors in the fingers. Figure 3A and 3B show the encounter regime and the proximity sensor display respectively with the hand misaligned. Figures 4A

and 4B show the encounter regime and the proximity sensor display with the hand aligned. In these figures the hand is shown schematically together with four bars indicating the d_i's of Figure 2. The bar lengths are proportional to the d_i's. At the bottom of the display the required corrective control is shown. In Figure 3 a large approach angle (yaw) error is shown. In Figure 4 that error has been eliminated. The error is much easier to see from the automatically monitored error bar under the A than it is from comparing the relative lengths of the d_i's visually or from examining the scene in a TV view.

2. Event-Controlled Displays

Changes in sensor data can automatically effect changes in both display formats and display parameters thus matching the particular information required for manipulation motion control to different phases of the task. Consider the task where the operator wants to move a manipulator from a safe uncluttered area to one with a rich environment, locate and pick up an object, and remove it to an uncluttered area. To perform this task the operator will need general TV views, detailed proximity information, and a knowledge of the forces and torques. Contact pressures and areas may also be useful.

Throughout the manipulative task many different displays and display modes are required. The display requirements during a gross motion phase of a task are not the same as those in the encounter regime. Similarly even within the encounter regime the operator may want proximity data combined in a single view with touch or other sensor data merged or separate from any TV camera views that are available. The event controlled display technology described here have been implemented using the first three data types described above, shown separately from the TV scene data.

Photographs of the three displays employed are presented in Figures 5, 6, and 7. The proximity sensor scene displayed corresponds to the case where the hand is level over, say, a table top and no object is in front of the fingers. The force/torque displays show both the quantitative forces and torques, in ounces and inch-ounces respectively, as well as graphical bar representations of the data. For each bar zero is in the center of the screen, positive data values generate a bar to the right, and negative values a bar to the left. In the combined, "dual", display both proximity and force/torque data are displayed, although in reduced scale.

A transition diagram for the six subtasks required to perform this scenario are given in Figure 8. At (1) the manipulator is in the safe condition. The proximity only display is enabled to focus his attention on getting into the encounter regime and aligning the manipulator to the object without accidental collision. When alignment has occurred, (2), the displays should show both proximity and forces and torque. Together with the TV display this allows the object to be touched without large unknown forces and while maintaining alignment. At this point, (3), the operator removes the biases in the force/torque readings caused by cables and gravity and initiates the grasp process. As the object is grasped, knowledge of the forces and torques becomes of paramount importance and the force/torque only display mode is entered, (4). As the object is lifted its contact with environment is reduced, (5), and once

again the combined proximity and force/torque display is desired. As the object is moved into a still less cluttered environment, the hand biases may be restored as they are manipulator geometry and gravity vector dependent so that the operator has a full view of the manipulator loads. The manipulator is then brought into the safe region, (6). Note that if the object is to be placed in a new rich environment steps (2) to (6) may be repeated several times. A substantial workload is placed on the operator to manage these display mode transfers particularly when he may already be using his hands to control the manipulator through supervisory computer control or through dual joy stick control. The problems are compounded if the operator is employing two manipulators to perform a task.

Automatic event mode switching can alleviate much of the display control workload. The conditions for steps (2), (4), and (5) can be detected using event detection logic and the mode switches automated. This scheme has been implemented and is undergoing testing and evaluation. This automatic mode switching is expected to reduce the operator's workload whether the manipulator is directly controlled in a manual mode or whether it is controlled in a supervisory mode.

3. Plans

Automatic display mode switching is still in an initial phase of development. Subsequent work is planned, first, to refine the definitions of "display state transitions" or "display mode switchings." A fundamental concern is to guarantee stability and continuity in the human perception of the overall control task. Second, to conduct performance tests emphasizing the human response to the automatic display mode switchings during the test evaluation. The plans also include the design and implementation of new display formats which can simplify the perception of complex multivariable error states.

III. VOICE CONTROL OF DISPLAYS

The development of voice controlled displays, displays whose parameters, format, data base, event detection logic, or state transition matrix can be changed by the operator's voice commands, alleviates a major bottleneck in the man-machine interface: the problem of his hands bounding the manipulator man-machine communication channel bandwidth, and enables the man to perform more functions in a simultaneous, coordinated fashion.

Voice controlled displays can be applied to either manual or computer control modes. In either mode they enable the operator to focus on the principal displays and control functions and to monitor task progress without interruption. To change display modes the operator is not required to remove his hands from, say, either the joint angle joy sticks or a computer console.

In most applications the commands voiced by the operator do not have to be verified as they would for voice control of the arm itself. This distinction occurs because the changes to the displays are in the information

feedback, not in the task execution feed forward where changes can cause the task to be done incorrectly. To recover for a mis-voiced display command the operator need only voice new commands to cycle back through the display transition matrix to the desired state.

1. Implementation

Figure 9 shows the voice control display transition matrix implemented. As before, the proximity, dual, and force/torque circles shown denote display modes. In addition, a monitor mode has been added. This monitor function globally restores all the display parameters and; transition sequences to a known state, include the disabling of event modes. The transition between states can be made by the operators voicing the desired destination state, e.g., "proximity", "both" (for dual), "force" (for force/torque), or "monitor". In the proximity or force/torque states the operator may also enter commands by voice or keyboard to control event parameters (magnitude, tolerance, status (on/off), or effect).

The syntax diagram is given in Figure 10. The syntax contains two primary control words, "camera" and "display" which constrain the word list over which a search for a match is made. In this way a higher word recognition reliability can be maintained, and a faster match can be achieved. The keyboard and voice display commands are shown in Table 1. A cursory view of the computer network employed is shown in Figure 11. The network computers each are dedicated to specific functions. The Interdata microcomputer to arm commands and control, the Nova microcomputer (in an Interstate Electronics discrete word recognition system) to voice message generation, and the Z-80 microcomputer to real time displays (Refs. 6 and 7). Note that the loop for control of the TV cameras is included for completeness, although it is not yet implemented, except in the discrete word recognition command vocabulary.

Thus far the voice command of displays has been tested to verify that the voice command system works reliably in a noisy laboratory environment, to discover where changes to the vocabulary should be made because words have too similar a sound, and to begin to evaluate the presently implemented display commands. While most words are recognized adequately (>95% accuracy with training) in the laboratory environment, some words are confused and modifications will have to be made to the command word list. For example, while the words "force" and "first" are natural and task related they are difficult for the voice system to reliably separate. Thus one or the other should be modified to a new natural word. Similarly, while most of the display commands, formats and state changes are useful, some additions are needed to give the operator the ability to change the parameter sets in a more task related manner.

2. Plans

The plans include the conduct of experiments using alternative command vocabularies and the increase of the variety of display modes which can be controlled by voice commands. There are two basic concerns: (a) the voice recognition reliability and (b) the naturalness of command words and syntax.

TABLE 1

KEYBOARD AND VOICE DISPLAY COMMANDS

Word(s)	Effect
Remove	Clear display. Transfers control to monitor
Proximity	Turns on proximity display
Force	Turns on force/torque display
Both	Turns on combined display
First	Selects first set of event parameters
Second	Selects second set of event parameters
Magnitude, nnn	Sets magnitude parameter for event mode*
Tolerance, nnn	Sets tolerance parameter for event mode*
Event	Enables event mode
Off	Disables event mode
Auto	Enables automated display transition
Bias	Biases force/torque data. Current data shown is 0
Clear	Clears bias. Real force/torque data shown
X,Y,Z,V	Enable X, Y, Z or Vector force event modes
1,2,3,4	Enable one of four proximity event modes
nnn	Decimal quantity. Oz units for force. 0.1 inch units for proximity

*Different parameter lists are kept for forces and proximity. Parameters can be entered from corresponding display mode.

It is also planned to implement the control of the TV cameras by discrete words commands.

IV. CONCLUSIONS

The higher level automation of sensor data display control and the voice control of displays open as yet unexplored avenues to improve the utility of displays for remote manipulator control. Performance tests are required to

evaluate the benefits of both developments in quantitative terms. However, the developments and experiments accomplished thus far suggest their value and areas for futher development. In particular, further developments are required to fully integrate event displays with TV displays of the visual work scene.

REFERENCES

1. Bejczy, A. K., Paine, G., "Event-Driven Displays for Manipulator Control," Proceedings of the 14th Annual Conference on Manual Control, U.S.C., Los Angeles, California, March 1978.

2. McGovern, D. E., "Factors Affecting Control Allocation for Augmented Remote Manipulation," Ph.D. Thesis, Stanford University, Stanford, California, 1974.

3. Hill, J. W., "Study to Design and Develop Remote Manipulator Systems," Annual Report 1, Contract NAS2-8652, Stanford Research Institute, Menlo Park, California, July 1976.

4. Bejczy, A. K., "Environment-Sensitive Manipulator Control," Proceedings of the 1974 IEEE Conference on Decision and Control, Phoenix, Arizona, November 1974.

5. Bejczy, A. K., "Smart Sensors for Smart Hands," AIAA Conference on Smart Sensors, Hampton, Virginia, November 1978, paper no. 78-1714.

6. Bejczy, A. K., Zawacki, R. L., "Computer Aided Manipulator Control," Proceedings of the First International Symposium on Mini- and Microcomputers in Control, San Diego, California, January 1979.

7. Paine, G., "Microprocessors for Realtime Displays and Control of Space Teleoperators," Proceedings of the First International Symposium on Mini- and Microcomputers in Control, San Diego, California, January 1979.

ACKNOWLEDGMENT

The contributions of Mr. R. L. Zawacki to this work, the programming of the display control vocabulary and design of the format for the data transfer from the voice recognition computer to the display control computer, are gratefully acknowledged.

The research described in this paper was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under NASA Contract NAS7-100.

APPENDIX A
SYSTEM DESCRIPTION

The Advanced Teleoperator Development Laboratory contains: two manipulator systems (a CURV parallelogram linkage slave system and an anthropomorphic master/slave system); a discrete word recognition system from Interstate Electronics containing a NOVA 2 computer; a VOTRAX speech synthesis unit; a display system driven by a Z-80 microprocessor capable of realtime black and white and color graphics displays; an Interdata Model 70 computer for arm control in either supervisory or manual control modes; and a multiple input control panel capable of position and rate inputs through potentiometers or joysticks. The laboratory occupies three separate rooms devoted to: 1) the manipulators and their hardware and software development; 2) a control station and voice input development; and 3) display and sensor development, respectively. The details of the laboratory are presented in 5 and 6. An overview is shown in Figure 11.

The three principal elements employed in the work described here are the discrete word recognition system, the arm based sensors, and the display system. During testing and evaluation the Interdata control computer, the manipulators, and the control hardware and software will be employed. The task scenarios, the operator procedures, the supervisory software, and the control modes are important in all phases of this work.

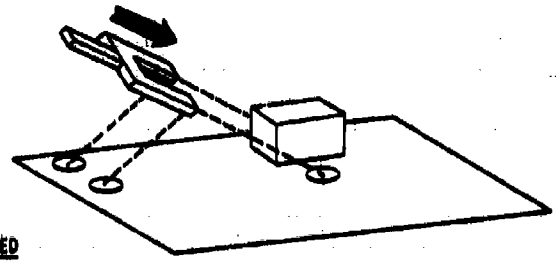
The VOTRAX speech recognition system is capable of storing patterns for several hundred discrete words. Each pattern consists of 240 bits of data related to the energy content versus time for a number of frequency bands. To reduce the number of word patterns that must be matched, both speeding up the match process as well as increasing its accuracy, the number of permissible words can be limited at any one time. Presently for either manipulator or display control the operator is constrained to know and follow a prescribed syntax to communicate his message. For example, the sequence "display", "proximity" is valid and would cause the proximity display to appear where "display", "camera", "proximity" would leave the VOTRAX in a syntax state expecting further camera related inputs and would have no effect on the display system.

The proximity sensors measure the intensity of the light reflected by an object placed in front of them. They contain a matched light-emitting diode and photodetector plus associated optics and optical filters. New designs couple the electro-optical components to the manipulator hand with fiber optic bundles simplifying the system design and improving its performance. To reject ambient light the emitted light is amplitude modulated and a synchronous detector employed. If the sensed object's surface parameters (spectral and spatial reflectivity) are known or can be controlled, accuracies of 0.3 cm or less are possible for a sensor with a range of 15.0 cm following calibration.

The force-torque sensors were manufactured by Vicarm, now with Unimation, and contain 16 semiconductor strain gauges mounted on a cruciform bridge. These gauges are in pairs to reduce thermal effects. The eight resultant

readings can be transformed to sensed forces and torques through a 6x8 calibration matrix. Using the 16 most significant of the 48 elements, an accuracy of 5% can be achieved. The useful range of the sensors is 0.5 to 300 N.

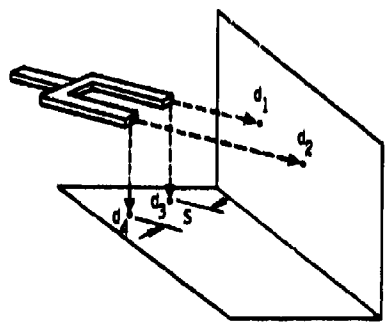
The display system consists of a Z-80 microprocessor based computer for software development and display driving, color and black and white TV interfaces, and interfaces to the anthropomorphic and CURV arms. The interface to the CURV arm provides all of the feedback signals that are available to the Interdata control computer, i.e., proximity and force/torque data, arm joint angles, finger separation, grip force, etc. The computer proper is built of standard S-100 bus components and has 56 k bytes of memory; dual floppy disks; and four serial, 10 parallel and 7 analog I/O channels.



INFORMATION NEEDED

- GEOMETRY (RANGES, ANGLES)
- RATES
- OBJECT PARAMETERS
- FINGER SEPARATION

Figure 1. Encounter Regime



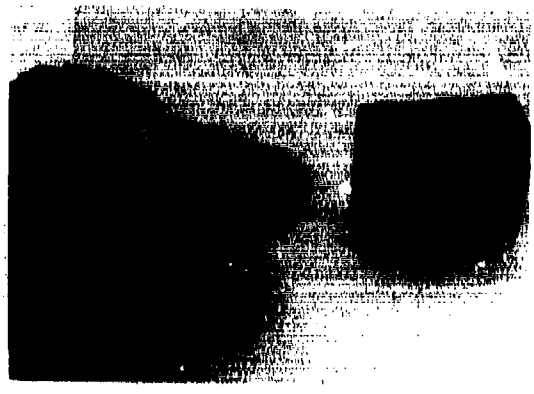
TARGET RANGE: $\frac{d_1 + d_2}{2}$

HEIGHT: $\frac{d_3 + d_4}{2}$

APPROACH, ANGLE: $\text{TAN}^{-1} \frac{d_1 - d_2}{S}$

HOVER ANGLE: $\text{TAN}^{-1} \frac{d_3 - d_4}{S}$

Figure 2. Encounter Geometry



A. APPROACH SCENARIO



B. PROXIMITY DISPLAY

Figure 3. Misaligned Encounter



A. APPROACH SCENARIO



B. PROXIMITY DISPLAY

Figure 4. Aligned Encounter



Figure 5. Proximity Display

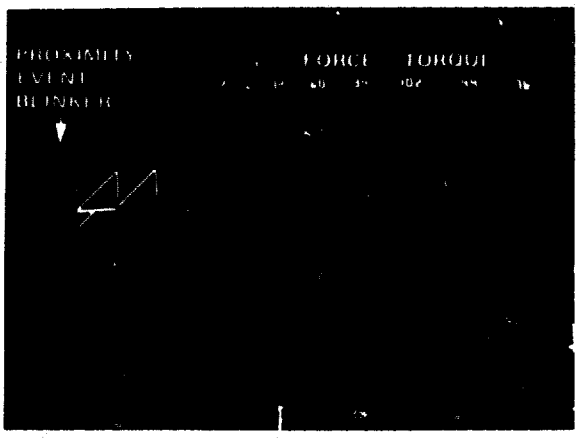


Figure 6. Combined Display

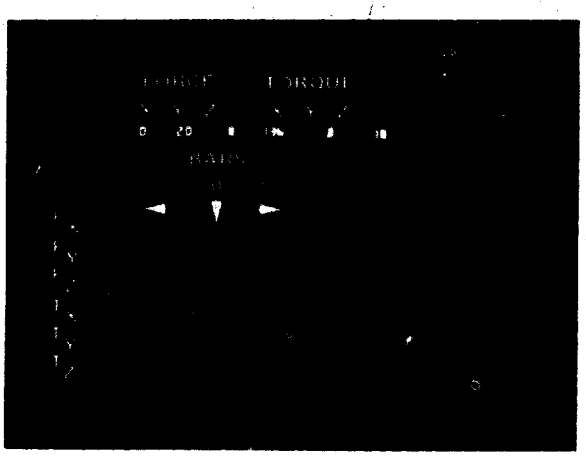


Figure 7. Force Torque Display

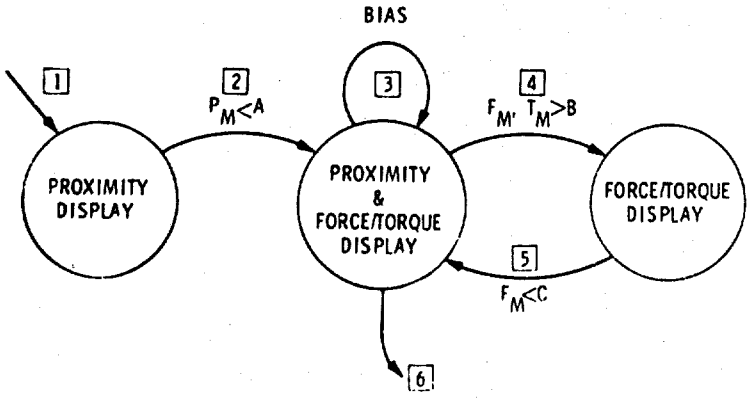


Figure 8. Event Controlled Display for Manipulator Approach, Load Acquisition and Removal

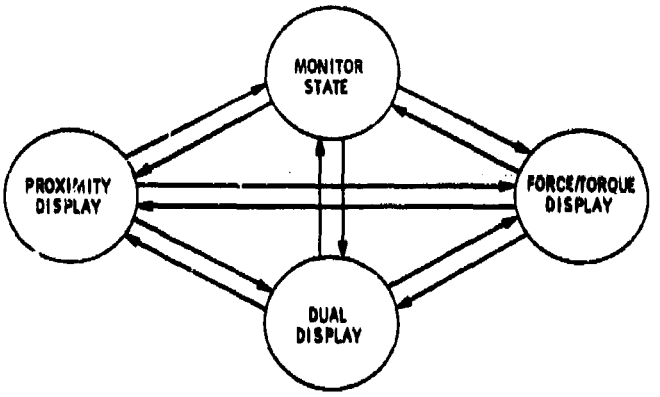


Figure 9. Voice Controlled Display Transition Diagram

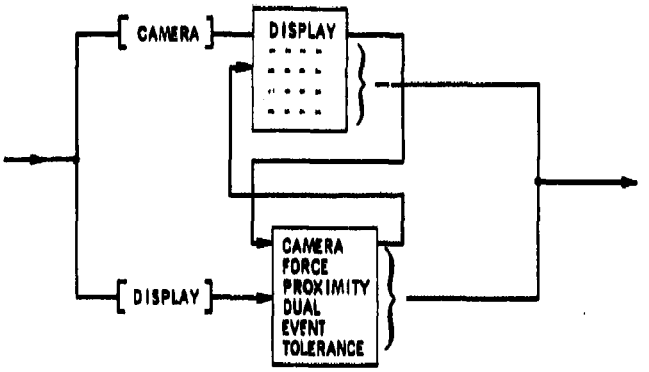


Figure 10. Voice Control Syntax

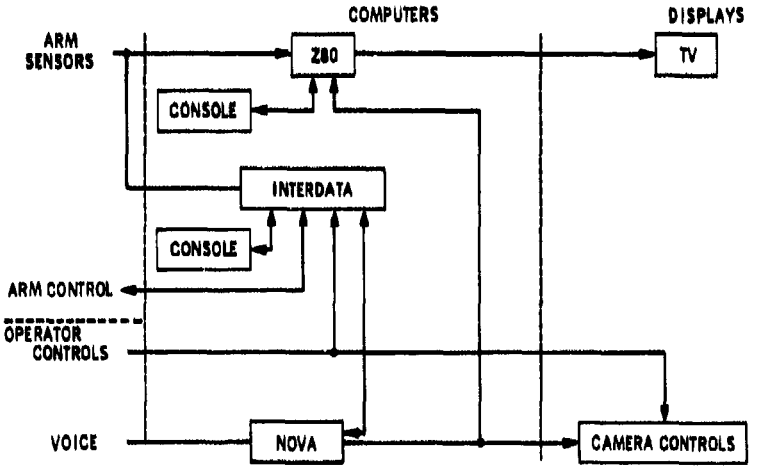


Figure 11. Voice Control Computer Net

FURTHER CONSIDERATION ON EXTENDED SIGNAL QUICKENING FOR MANUAL CONTROL

Masayoshi Tomizuka, Assistant Professor, and
Shigeru Kurosu, Research Associate
Department of Mechanical Engineering
University of California
Berkeley, California 94720

ABSTRACT

This work is concerned with extended quickening, a technique to facilitate human control over higher order plants with weak damping. In this technique the signals displayed are the quickened plant output and the quickened reference input. The quickened output is a weighted sum of the plant output and its derivatives, the role of which is to stabilize the feedback loop. The quickened reference input is a weighted sum of values of future reference input, which makes it possible to achieve high quality tracking. The technique is examined and evaluated under the following three conditions: 1) imperfect measurement of plant state variables (estimation of plant output derivatives), 2) presence of input noise, and 3) presence of additional time delay. Experiments conducted with a triple integrator plant confirmed that the technique is still effective under these nonideal conditions.

INTRODUCTION

Feedback controllers for higher order undamped systems must include first and higher order derivative actions to achieve the closed loop stability. In manual control, it is extremely difficult for the human operator to visually extract second and/or higher order derivatives of the displayed signal. Therefore, most human operators cannot control third or higher order plants with little or no damping with conventional compensatory or pursuir display [1]. To overcome this problem, Birmingham and Taylor [2] proposed the signal quickening technique in which a weighted sum of the plant output and its derivatives (a quickened output) is displayed in place of the output itself. The technique has been extended to incorporate future reference trajectory information to achieve a high quality tracking [3]. In the extended quickening in [3], future trajectory information is processed by a computer to generate a quickened reference which becomes the target for the quickened output under human control. A method for determining extended quickening parameters has been established based on discrete optimal control theory, and effectiveness of the technique for a triple integrator plant has been experimentally confirmed [3].

This paper further examines the extended quickening technique from a realistic point of view. Namely, we will consider the following points which were ignored in the previous study:

a, Estimation of derivatives of the plant output required for

constructing the quickened output,

b. effects of plant input noise; its location and strength,

c. effects of a pure time delay in the plant.

OVERVIEW OF EXTENDED QUICKENING

Figure 1 shows the overall picture of the extended quickening experiment discussed in this paper.* As in the previous paper [3], a triple integrator is implemented on an analog computer. An LSI-11 microcomputer is used for generating reference trajectories, estimating first and second order derivatives of the plant output, computing extended quickened signals (quickened plant output and quickened reference trajectory), realizing a pure time delay portion of the controlled plant and on-line data acquisition. The extended quickened signals are displayed on a CRT with a 12.7 cm diameter. The quickened reference signal is displayed by a dot in the center of the screen and its travelling distance on the screen is approximately ± 0.5 cm. The quickened plant output is displayed by two dots placed both sides of the quickened reference. The distance between subject and display is about 50 cm. The human operator moves the joystick so that the quickened output follows the quickened reference trajectory. The signal from the joystick is delayed through shift registers in computer software and then is applied to the triple integrator. The reference trajectory is a Gauss-Markov random signal which was generated

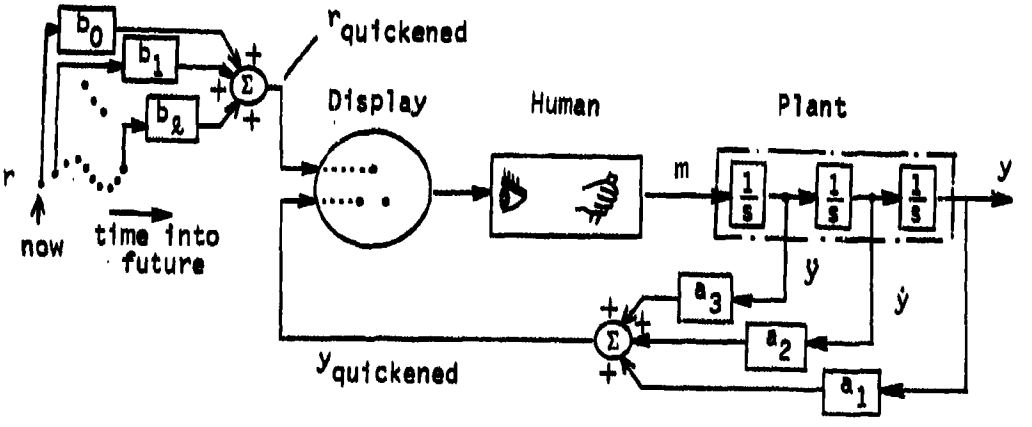


Fig.1 Extended Quickening

* a_i's and b_j's in Figure 1 are the extended quickening parameters. They are determined by the optimal control procedure described in [3]. The sampling period Δt in this paper is 0.025 sec (same as in [3]).

by a second order digital filter excited by a Gaussian white signal. An approximate bandwidth of the reference signal is 1.5 rad/sec. Each experimental run comprised a 100 second performance test preceded by a 20 second warming up period. The subject had an ample amount of training before the data was taken.

STATE ESTIMATOR AND INPUT DISTURBANCE IN EXTENDED QUICKENING

Computation of the quickened output requires the derivatives of the output. In [3], the derivatives were directly measured on the analog computer. However, this kind of direct measurement is usually not possible. In such cases, the missing derivatives must be estimated from the plant input and output data. Beside the state estimation problem, input noise or disturbance is unavoidable in most practical situations. In this section, the extended quickening is investigated under realistic conditions that an input noise is applied to the plant and that a Kalman filter is used as a state estimator. We do not consider the pure time delay in this section.

The whole scheme is depicted in Figure 2. A random disturbance is added to the controlling input to the triple integrator plant. It is a Gaussian white signal with variance W. The controlling input of human operator and the plant output are assumed to be the only measurable quantities and they are fed to the LSI-11 computer. The quantization error of the A/D converter is considered as measurement noise with covariance V.

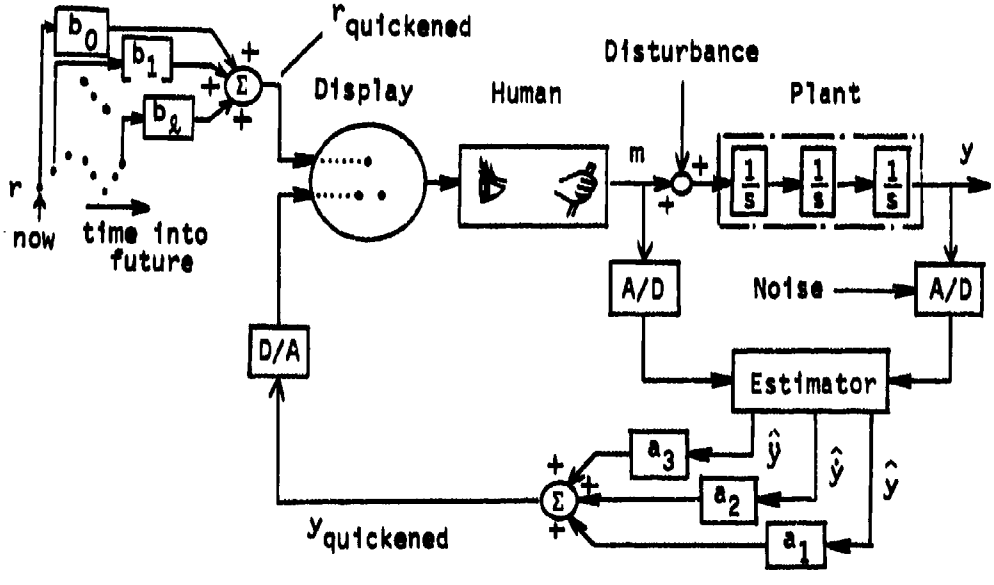


Fig.2 Extended Quickening with State Estimator

Estimator Design

The plant equation and output equation are given by

$$\underline{x}_p(k+1) = \underline{A}_p' \underline{x}_p(k) + \underline{B}_p' m(k) + \underline{B}_p' w(k) \tag{1}$$

$$y(k) = \underline{C} \underline{x}_p(k) + v(k) \tag{2}$$

where $\underline{A}_p'$ and $\underline{B}_p'$ are the same matrices as in [3], $\underline{C} = [1\ 0\ 0]$ and $w(k)$ and $v(k)$ are input and measurement noise respectively. To simplify the estimator design, it is assumed that $w(k)$ and $v(k)$ are both zero mean, white Gaussian noise processes and that they are uncorrelated. The state of the plant is then estimated by the discrete time Kalman filter

$$\hat{\underline{x}}_p(k+1|k) = \underline{A}_p' \hat{\underline{x}}_p(k|k) + \underline{B}_p' u(k) \tag{3}$$

$$\hat{\underline{x}}_p(k+1|k+1) = \hat{\underline{x}}_p(k+1|k) + \underline{F}[y(k+1) - \underline{C} \hat{\underline{x}}(k+1|k)] \tag{4}$$

$$\underline{F} = \underline{M} \underline{C}^T [\underline{C} \underline{M} \underline{C}^T + V]^{-1} \tag{5}$$

where $\hat{\underline{x}}_p(k+1|k)$ denotes the best estimate of $\underline{x}_p(k+1)$ based on the output data $Y_k \triangleq \{y(j)\}_{j=0}^{k}$, $\underline{F}$ is the Kalman filter gain and $\underline{M}$ is the steady state solution of the Riccati equation,

$$\underline{M}(k+1) = \underline{A}_p \underline{Z}(k) \underline{A}_p'^T + \underline{B}_p' W \underline{B}_p'^T \tag{6}$$

$$\underline{Z}(k) = \underline{M}(k) - \underline{M}(k)\underline{C}^T [\underline{C} \underline{M}(k) \underline{C}^T + V]^{-1} \underline{C} \underline{M}(k) \tag{7}$$

The covariance of the measurement noise, V, was selected to be 6.0×10^{-8} (volt^2) based on the standard deviation which was found from the A/D conversion of a very stable voltage.

The Kalman filter gain was calculated using this value of V and the given covariance of the input distrubance, W. Table 1 contains the Kalman filter gains for various values of W. Generally speaking, the larger the covariance of the input disturbance, the higher the Kalman filter gains. This causes a displayed output signal to be too shaky for W larger than 0.0001. The shaky displayed signal was found to be not helpful but confusing to the human operator.

W	f_1	f_2	f_3
10^{-8}	0.0556	0.0333	0.0099
10^{-6}	0.0875	0.1306	0.0975
10^{-4}	0.1631	0.5518	0.9337
10^{-2}	0.3117	2.2968	8.4672
ad hoc	0.35	2.5	8.0

V=6.0 10^{-8}

Table 1. Kalman Filter Gains

Effect of Filter Gain

The first set of experiments was conducted to examine the closed loop behavior with different combinations of feedback (or quickening) gains obtained with selected values of the design parameter w.*

Figure 3 shows the effects of the preview time upon the time average (RMS) tracking error and RMS controlling input. The data in the figure are normalized with respect to the RMS tracking error and controlling input for the zero preview case under the ideal consition [3]; i.e. direct measurement of output derivatives and no input noise. RMS values of the tracking error and controlling input are both improved as preview time is increased. The performance in term of RMS tracking error is better in the case with w = 0.01 than with w = 0.1. However, the smaller w is more demanding to the human in terms of controlling effort. We can see paradoxical results that the smaller the input noise covariance, the larger the RMS tracking error. This is due to the pole location of the Kalman filter (see Figure 4). We used the steady state Kalman filter as a state estimator and the gain was computed solely from this W and V. As can be found in Figure 4, the filter poles are asymptotic to the origin as W is decreased; namely, the steady state Kalman filter is asymptotic to the open loop estimator, the dynamics of which are identical to the limitedly stable plant dynamics. So it is not desirable to use a filter gain computed with a too small W. Therefore it was decided to use the estimator gain which gives the estimator poles labeled X in Figure 4 regardless of the value of input noise covariance. It was an 'ad hoc' approach, but worked in the experiment.

Figure 5 shows the effect of the preview time upon the RMS error and controlling input when the 'ad hoc' estimator gain was used. It is interesting that the input covariance does not have strong effect upon the tracking performance.

* See [3] for the definition of w and determination of gains.

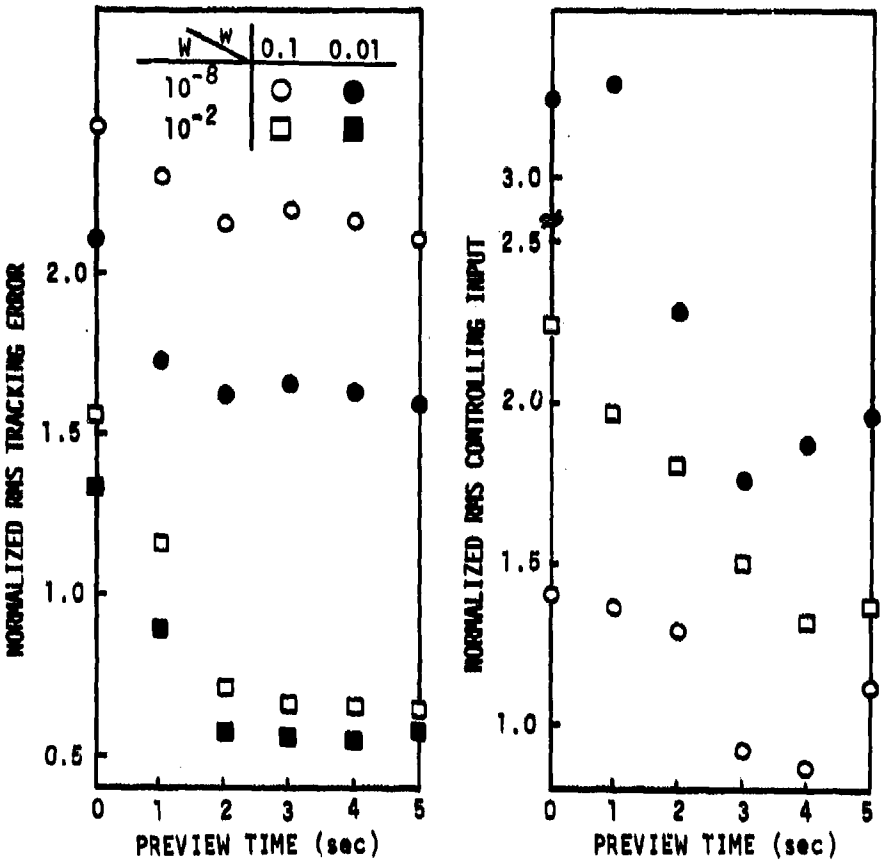


Fig.3 Effect of Input Disturbance

(Data Normalized with Respect to the Zero Preview Case under Perfect State Measurement and No Input Noise [3])

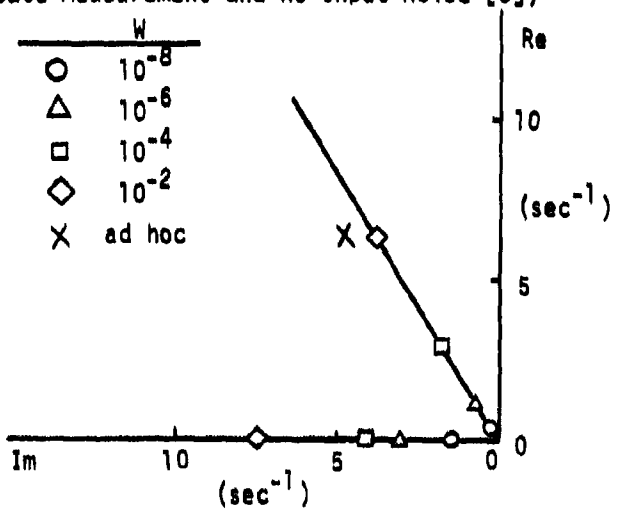


Fig.4 Location of Estimator Poles for Selected Values of Input Noise Covariances
(Poles in z-plane have been converted to these in s-plane)

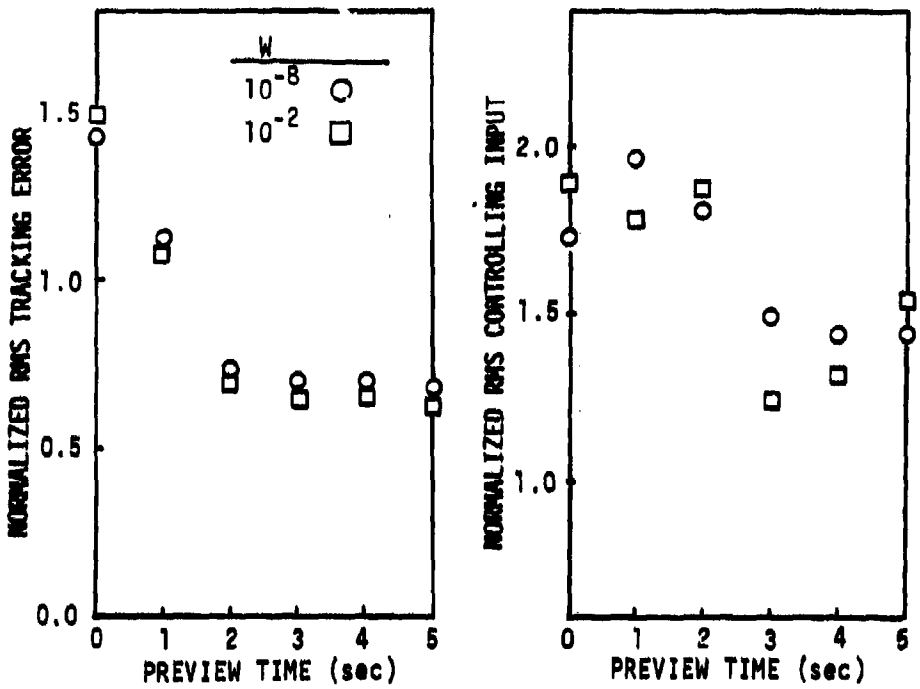


Fig.5 Effect of Input Disturbance with the 'ad hoc' Kalman Filter Gain
(Data Normalized with Respect to the Zero Preview Case under Perfect State Measurement and No Input Noise [3])

Effect of Noise Location

To investigate the effect of input noise location, experiments were conducted with input noise applied to the plant as shown in Figure 6. Again, the 'ad hoc' estimator gain was used in this set of experiments. Figure 7 shows the experimental data. It can be seen that the input noise applied at location 2 requires more contorlling effort of the human operator than the noise at location 1. This is due to the fact that the noise effect on the plant output is more evident when the noise location is closer to the output. Except for controlling input, the performance improvement by preview takes place essentially in the same pattern as before. This confirms that the 'ad hoc' estimator works for supplying the missing output derivatives.

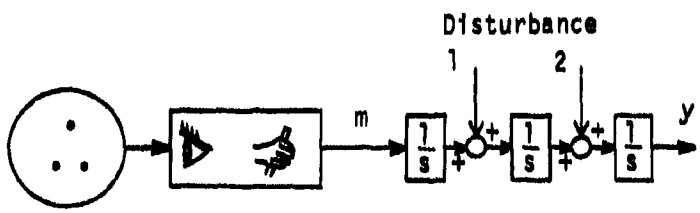


FIG.6 Input Noise Location

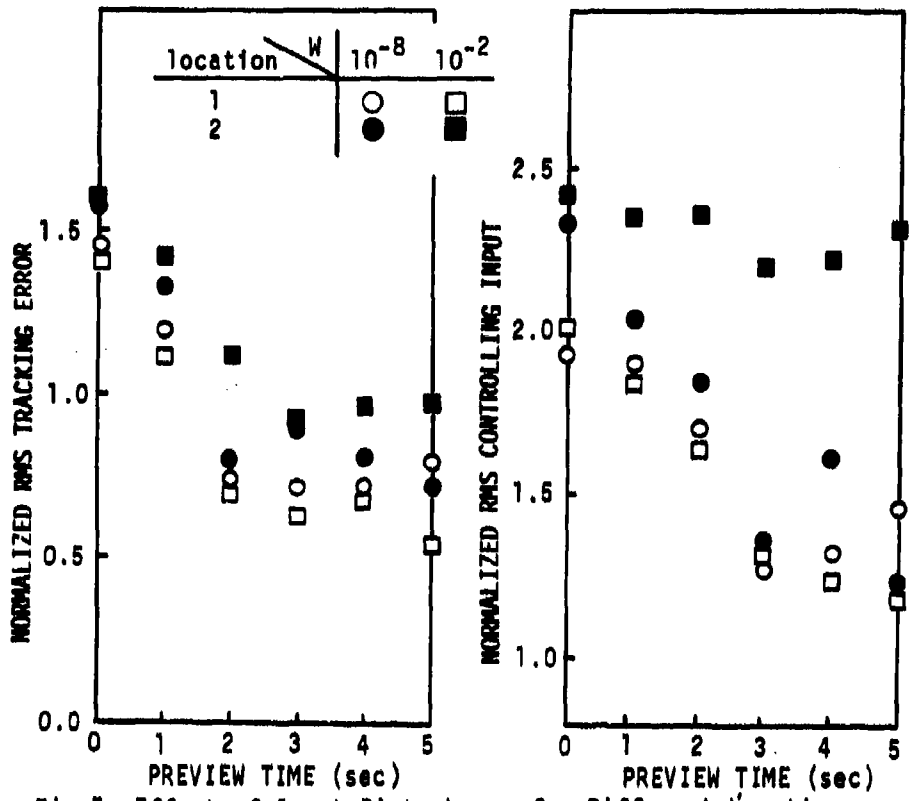


Fig.7 Effect of Input Disturbance for Different Locations (Data Normalized with Respect to the Zero Preview Case under Perfect State Measurement and No Input Noise [3])

INCLUSION OF A PURE TIME DELAY IN THE PLANT

To study the extended quickening in a little more complicated situation, a pure time delay is added to the triple integrator plant. In this section, the output derivatives are measured directly from the plant. Also, input noise is not considered.

Determination of Parameters in Extended Quickening

An added pure time delay τ_d increases the dimension of the discrete time model of the plant. The overall plant model is given by Eq (3) in [3] preceded by d_p-step time delay where $N_d = \tau_d/\Delta t$ and Δt is the sampling period. As in [3], the human operator is first modelled to be a pure time delay L, which corresponds to d-step delay z^{-d} ($d = L/\Delta t$). Therefore, the open loop human plant dynamics are given by Eq (3) in [3] preceded by $z^{-(d + N_d)}$, and the determination of the extended quickening parameters (i.e. feedback gains and

feed forward gains) can be done similarly as in [3]. It should be noted that the dimension of the matrix Riccati equation in the present case becomes $(3 + d_p + d) \times (3 + d_p + d)$.

The structure of extended quickening for the plant with a pure time delay, is depicted in Figure 8. The portion of the structure inside the dashed line can be viewed as human operator. The reason for this is the same as in [3].

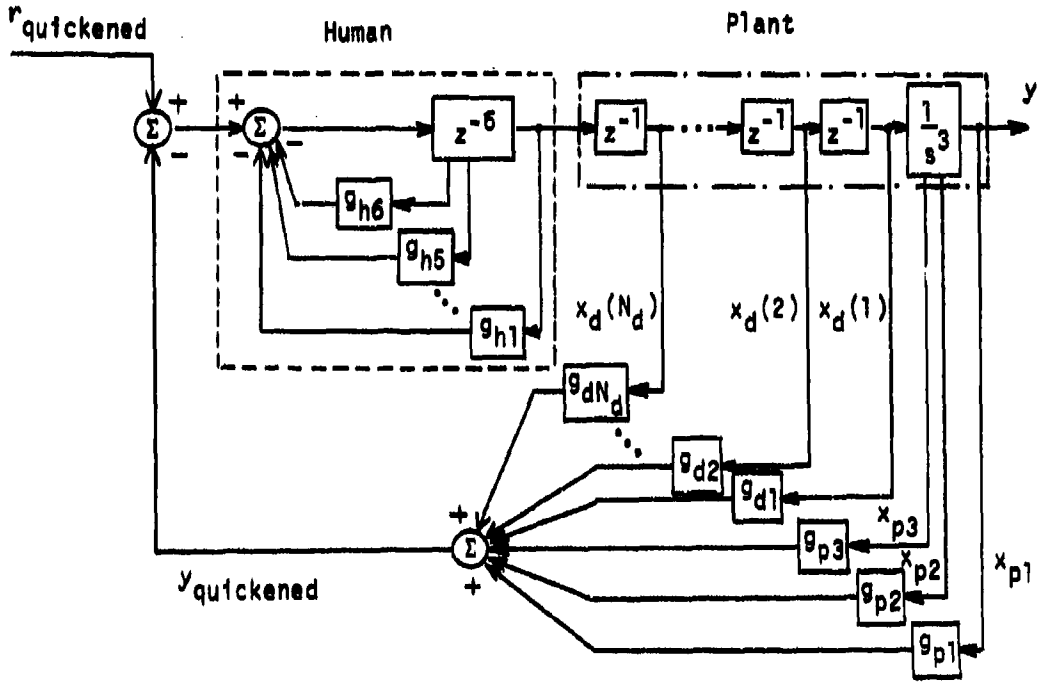


Fig.8 Structure of Extended Quickening Based on Optimal Control

For selected values of τ_d and w = 0.1, the steady state solution of the Riccati equation was computed and the feedback gains, g_{pi}'s (i = 1,2,3), g_{di}'s (i = 1,2,...,N_d) and g_{hj}'s (j = 1,2,...,6) were computed. Results are summarized in Figure 9. From the figure, we find that g_{pi}, g_{hi}'s are not affected by a pure time delay as compared to the no delay case [3] and that the feedback gains for output derivatives become larger as τ_d is increased. This is because a larger time delay requires a stronger damping for the closed loop stability.

Effect of Time Delay

An experiment was conducted to examine the effect of different sets of feedback gains obtained by varying the length of the time delay, and to verify the performance improvement that can be achieved by extended quickening. In

τ_d (sec)	g_{p1}	g_{p2}	g_{p3}	g_{h6}
0.0		0.160	0.168	
0.2		0.175	0.202	
0.4		0.190	0.238	
0.6	0.0757	0.205	0.278	-0.9129
0.8		0.220	0.320	
1.0		0.235	0.366	

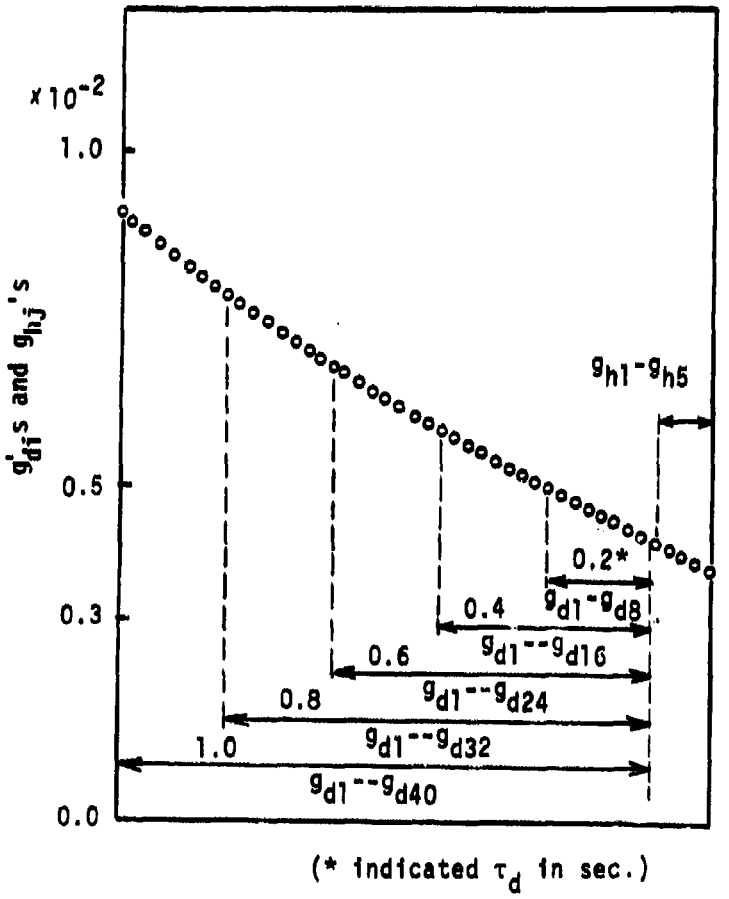


Fig.9 g_{p1}'s, g_{d1}'s and g_{h1}'s (w=0.1)

the experiment, the time delay was implemented by computer software: i.e. equations for N_d stage delay chain were included in the computer program.

Theoretically, all the stored data in the shift register must be fed back as well as plant output and its derivatives. From Figure 9, however, feedback gains for the delay chain are orders of magnitudes smaller than other feedback gains, g_{p1}, g_{p2} and g_{p3}. Hence the first set of experiments was conducted ignoring the feedback from the delay chain. Figure 10 indicates that the RMS tracking error and controlling input become larger as τ_d increases. It can also be found that the larger the time delay, the more the controlling effort. For τ_d = 0.6 sec, it is very difficult to control the plant. However, the performance improvement by preview is as expected.

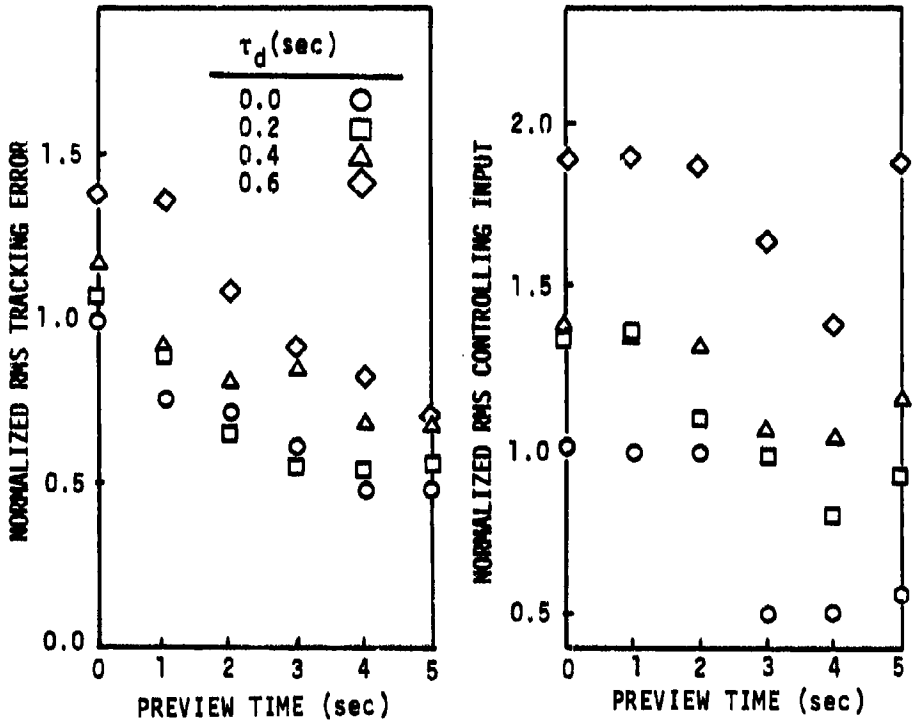


Fig.10 Effect of Time Delay -- No Feedback from Delay Chain (Data Normalized with Respect to the Zero Preview Case under Perfect State Measurement and Zero Delay Time)

In the final set of experiments, the feedback of state in the delay chain was included. Since it was not practical to feedback every state in the delay chain, we treated 8 delay elements as a group and used one feedback gain for each group. Namely, the quickened output is

$$sp(k) = \sum_{v=1}^{3} a_i \, x_{p1}(k) + \sum_{\xi=1}^{Nd/8} a'_d \, x_d[8(\xi-1)+1] \qquad (8)$$

where $a_1 = 1$, $a_2 = g_{p2}/g_{p1}$, $a_3 = g_{p3}/g_{p1}$,

$a'_d = a_{d8(\xi-1)+1} + a_{d8(\xi-1)+2} + \cdots + a_{d8\xi}$,

$a_{di} = g_{di}/g_{p1}$ and $x_d(i)$ is the i-th state in the delay chain as defined in Figure 8.

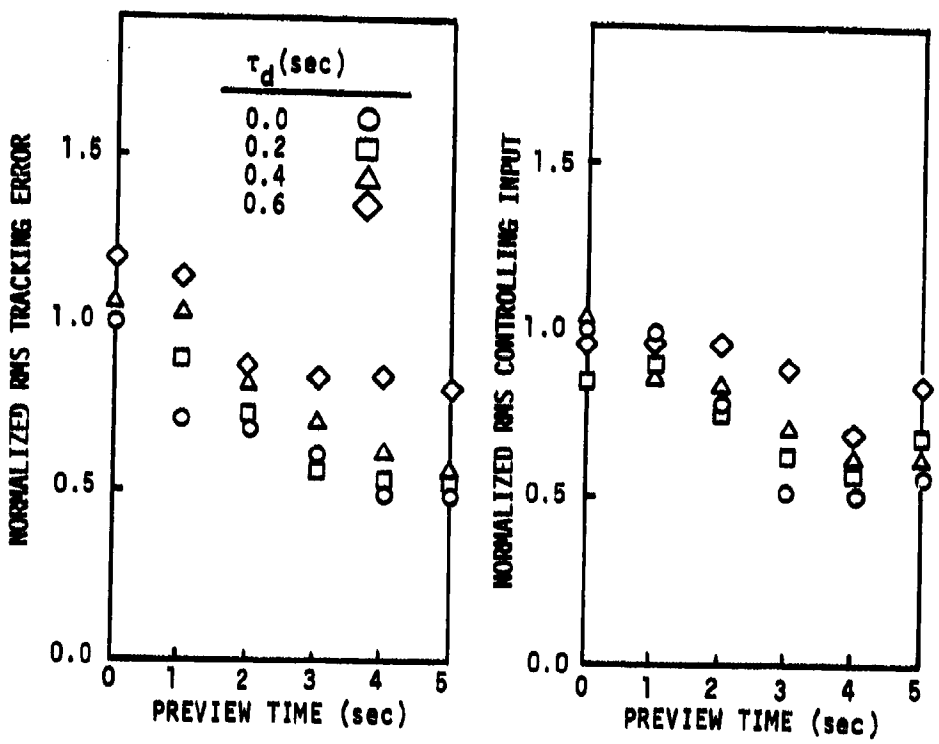


Fig.11 Effect of Time Delay -- with Feedback from Delay Chain (Data Normalized with Respect to the Zero Preview Case under Perfect State Measurement and Zero Delay Time)

Figure 11 shows the RMS data obtained by experiment. It can be seen that the performance improvement is again the same pattern as before. Notice that the RMS controlling input is much smaller relative to the case that the feedback loop from the delay chain was ignored. From these results, it was concluded that the extended quickening technique is effective even in the presence of a pure time delay in the plant.

CONCLUSION

The extended signal quickening was investigated from a practical viewpoint which included:

1. Estimation of derivatives of the plant output required for constructing the quickened output,

2. effects of plant input noise: its location and strength,

3. effects of a pure time delay in the plant.

The extended quickening technique was found to be effective under the influence of input noise and imperfect state measurement. The Kalman filter with an 'ad hoc' gain was a good estimator for supplying and missing state variables under a variety of conditions: different noise strength and locations.

The technique was also proved to be effective for the plant with a pure time delay. As predicted by the theory, the experiments indicated that feedback from the delay chain should be included for achieving a good performance. The work reported in this paper is being continued to validate several assumptions used in the study. One interesting question is whether a first order lag with a pure time delay (Eq (15) in [3]) is a good model of the human operator in the extended quickening.

REFERENCES

1. T. B. Sheridan and W. R. Ferrell, Man-Machine Systems: Information Control, and Decision Models of Human Performance, Cambridge, Mass., MIT Press, 1974.

2. H. P. Birmingham and F. V. Taylor, "A Human Engineering Approach to the Design of Man-Operated Continuous Control Systems," Naval Res. Lab. Rep. 4333, April 1954.

3. M. Tomizuka and W. M. Tam, "An Extension on the Quickened Display for Manual Control," Proceedings of the 15th Annual Conference on Manual Control, Nov. 1978.

VISUAL SCENE PERCEPTION - FREQUENCY DOMAIN DATA AND MODEL PARAMETER ESTIMATION PROCEDURE

by P.H. Wewerinke
National Aerospace Laboratory NLR
the Netherlands

ABSTRACT

Results of a theoretical study and an accompanying experiment are presented concerning the visual scene perception process and its impact on aircraft approach performance.

A variety of visual scene characteristics was represented by the configurations involved. Their effect on approach performance and human information processing (primarily in terms of observation noise spectra and perceptual thresholds) was investigated utilizing a model of the visual scene perception process. The model is based on linear perspective geometrical cues and on relative (visual) motion cues and is combined with the optimal control model.

In addition, results are presented of an automated model matching procedure. Via a first-order gradient method a matching criterion (involving variances and frequency domain data) was minimized yielding a unique set of observation noise model parameter values. These values led to the perceptual thresholds associated with the visual cues involved.

INTRODUCTION

This study deals with the visual scene perception process and its impact on the visual approach performance. This process is described (modelled) on the basis of the linear perspective geometry and cues related to the relative position and movement of the observer with respect to the outside world.

After linearization of the mathematical relationships between the visual cues and the aircraft state variables the model can be integrated in the existing framework describing pilot-aircraft behavior (optimal control model).

A variety of visual scene characteristics have been represented by the configurations involved in this study to validate the visual scene perception model.

Specifically this paper describes a model matching procedure to derive from the experimental data the "best" estimate of the observation noise parameters of the model. These key model parameters are obtained by minimizing a matching criterion involving variances, pilot describing

function data and observation noise spectral density data via a first-order gradient method. Finally, these values will lead to the perceptual threshold values associated with the visual cues involved.

In addition, the experimental and model results will be discussed and related to the visual scene characteristics incorporated in the various configurations.

VISUAL SCENE PERCEPTION

The analysis of the visual scene perception process is extensively described in references 1 and 2. A summary of this analysis will be presented in this chapter yielding the background information for the following chapters.

Based on a concise inventory of the most important characteristics (cues) of the visual scene the visual scene perception process has been described (modelled) on the basis of the linear perspective geometry and cues related to the relative movement of the observer with respect to the outside world. This amounts to mathematical relationships between these visual cues and the aircraft state variables. After linearization this model could be integrated in the existing framework describing pilot-aircraft behavior (the optimal control model).

Ten configurations were selected representing various outside viewing conditions and control modes (vertical lateral and two-axis control). The configurations are summarized in table 1.

Model predictions were based on the assumptions concerning perceptual thresholds of the various cues derived from the visual approach scene (shown in figure 1), noise levels associated with observing these cues and the interference among them. Values for these parameters were derived from baseline experimental data supplemented by the psychophysical literature. Based on these values a theoretical analysis was performed resulting in model predictions of system scores and frequency domain measures (pilot describing functions and observation noise spectra).

In addition, an experimental program was conducted to validate the visual scene perception model. The results of this experiment, involving the above-mentioned configurations, were compared with the model predictions providing a test of the hypotheses underlying the model (or: allowing a verification of the values assumed for the model parameters).

The experimental results in terms of mean-squared system variables agreed relatively well with the model predictions[M] (see refs. 1 and 2) showing the predictive capability of the model. However, a more detailed comparison of the frequency domain measures exhibited relatively large differences primarily in the observation noise spectra. For this reason an automated model

[M] Once the assumption was made that there is no attentional interference between the vertical and the lateral task (see chapter 3).

matching procedure was followed yielding the "best" estimate of the observation noise parameters of the model, given the experimental data. This is discussed in the next chapter.

MODEL PARAMETER ESTIMATION PROCEDURE

In order to obtain a better agreement between the experimental and model results in terms of the frequency domain measures, the model observation noise covariances were optimized (yielding a minimal matching criterion). Only this model parameter was considered as a (dependent) variable. This was based on a sensitivity analysis and previous modeling experience showing that the observation noise covariances are the key model variables especially in this specific study of the visual perception process.

The remaining model parameters which were kept constant throughout the analysis were: a perceptual time delay of 0.2 sec, a negligibly small motor noise ratio and a neuromotor time constant of 0.25 sec (corresponding to outer loop control).

Parameter estimation procedure

The matching criterion involved a weighed sum of (M) mean-squared (MS) system variables, of the pilot describing function gains (G) and phase angles (P) of the (N) frequency points, and of the observation noise spectral density values (O) of the (N) frequency points:

$$L_t(V_{y_i}) = L_{ms} + L_{yp} + L_{on} \tag{1}$$

with

$$L_{ms} = \sum_{i=1}^{M} a_{ms_i} \left| \frac{MS_{i_m} - \overline{MS}_i}{\sigma_{MS_i}} \right| \tag{2a}$$

$$L_{yp} = \frac{a_{yp}}{2N} \sum_{j=1}^{N} \left| \frac{G_{j_m} - \overline{G}_j}{\sigma_{G_j}} \right| + \left| \frac{P_{j_m} - \overline{P}_j}{\sigma_{P_j}} \right| \tag{2b}$$

$$L_{on} = \frac{a_{on}}{N} \sum_{j=1}^{N} \left| \frac{O_{j_m} - \overline{O}_j}{\sigma_{O_j}} \right| \tag{2c}$$

where a_i represents a weighting factor; the subscript m refers to the model quantity related to the mean ($\bar{(\cdot)}$) and standard deviation ($\sigma_{(\cdot)}$) of the corresponding measured quantity. So the matching criterion involved a weighed average discrepancy in mean-squared system variables, the describing function gains (in dB) and phase angles (in deg) and observation noise spectral densities. Each quantity was weighed by the standard deviation of the measurements thereby taking into account the reliability of the measurements and providing the necessary normalization yielding non-dimensional units.

The weightings (given in table 2) were selected on the basis of a preliminary sensitivity analysis and of subjective judgement: most weight is placed upon the system outputs (approach angle α and centerline inclination angle ω_c), half of that weight upon the aircraft attitude angles (pitch θ and roll φ) and one fourth of that weight upon the control inputs (elevator δ_e and ailerons δ_a). The total weight of about one preserves the distinct meaning of the resulting L_{ms}: the matching error in the mean-squared values in units of weighed averaged number of standard deviations. The describing function error and the observation noise error are weighed equally. Also the total error in the mean-squared scores and the total error in the frequency domain measures is weighed equally.

Now, this matching criterion was minimized via a first-order gradient method (Refs. 3 and 4). The gradient algorithm involved an adaptive scheme for the observation noise covariance steps (ΔV_{y_i}) requiring typically 4 to 8 iterations to drive L_t within 1 percent of its minimum value. The result was the optimal model parameters V_{y_i}.

Once V_{y_i} has been found it can be related to the corresponding perceptual threshold according to (e.g. ref.5)

$$V_{y_i} = \frac{P_o \sigma_{y_i}^2}{f_i K_i^2(a_i, \sigma_{y_i})} \tag{3}$$

where P_o is the (relatively constant) "noise-signal" ratio, $\sigma_{y_i}^2$ is the signal variance, f_i is the fraction of attention dedicated to variable y_i and K_i is the describing function gain associated with a threshold (a_i). Assuming an optimal allocation of attention among the visual cues (yielding the minimal cost functional J (Ref.4)) f_i can be computed once V_{y_i} is known (Refs. 1, 4 and 6). Then, for an assumed level of attention P_o and given (measured) signal variance, the describing function gain K_i follows from eq(3). From K_i the threshold a_i can be computed numerically involving the inverted error function computation.

Model matching results

The model matching results are summarized in table 3 for the vertical control tasks and in table 4 for the lateral control tasks.

For the vertical tasks the scores are accurately matched (on the average within 0.1 standard deviation). The frequency domain match is, on the average, 0.8 standard deviation ($\omega_{y_p} + L_{on}$), which is primarily due to the describing function discrepancy. However, also considering the reliability of the data (Ref.1), the agreement between experimental and model results is relatively good (as will be visualized in the frequency domain plots given in the next chapter).

The resulting threshold values are also given in table 3. The values are obtained on the basis of three assumptions:

- all vertical and lateral tasks are performed with a constant level of attention P_o = -21 dB
- There is full interference between the various visual cues involved within one task; however, due to the favourable viewing characteristics of the visual scene there is no interference between the vertical and lateral task (Ref.1)
- attention which has to be divided among the various visual cues is optimally allocated.

From the threshold values corresponding with a specific viewing condition an average estimate is derived (α, $\dot{\alpha}$ and $\bar{\theta} = \theta - \alpha$). With the resulting perceptual thresholds values the matching criterion was computed again (table 3b). Only the system scores are now somewhat less well matched. The frequency domain match is unaffected. These model results will be compared with the experimental measures in the next chapter, containing also a discussion of the visual approach scene results.

For the lateral and roll (conf.7) control tasks the mean-squared scores are matched, on the average, within 0.2 standard deviation. The frequency domain match is, on the average, 1.0 standard deviation. The same procedure has been followed as for the vertical tasks to arrive at the ultimate perceptual threshold estimates. With these threshold values the final matching results were obtained (table 4b). Again, only the match of the mean-squared scores is affected and deteriorated to 0.65 standard deviation. Of course, a still better match could be obtained when abandoning the model parsimonious approach to match the data of all configurations by varying only the thresholds a_i.

EXPERIMENTAL AND MODEL RESULTS

Comparison of experimental and model results

The mean-squared system variables of all configurations are summarized in table 5. A comparison between the experimental and theoretical scores is facilitated by figure 2 showing also the variability of the data.

On the average, the scores differ by about 10 % which value can be considered to be quite small. Also in terms of the variability of the data (Fig.2) this difference (on the average 0.3 standard deviation for the vertical tasks and, on the average, 0.6 standard deviation for the lateral tasks) is considered to be statistically insignificant.

Measured and model pilot describing functions and observation noise spectra are compared in figures 3-6 for configurations 1, 2, 4 and 5. The frequency domain measures of the remaining vertical and lateral configurations were almost identical to confs. 2 and 5, respectively, and are not given. This result will be discussed in the next section.

In general, the agreement between the experimental frequency domain measures and the measures obtained from the model is quite good in the mid-frequency region (where "most of the action" is). This may be expected when all the mean-squared scores are closely matched. On the average, the difference in the frequency domain data of all configurations is within one standard deviation.

Discussion of the visual approach scene results

In this section the experimental and model results will be considered and related to the visual scene characteristics incorporated in the various configurations.

The effect of the runway visual range on the vertical control performance can be appreciated by comparing the results of conf. 1 (only the inclination of the runway sides) and 2 (providing the runway depression angle). Figure 7a shows that a significantly[M] superior approach performance (in terms of the approach angle α) is obtained when the runway depression angle is visible. This is accompanied by a substantially smaller pitch angle and control activity. The improved approach performance is also reflected by the higher pilot describing function gain shown in figure 8. Especially the observation noise spectra which are also given in this figure provide additional insight in the effect of the visual cues involved. The effect of the runway depression angle is the most apparent in the relatively high frequency region, i.e. related to the approach angle rate ($\dot{\alpha}$) information. This is also reflected by the perceptual threshold values resulting from the model matching procedure given in table 6. The value of 0.23 deg/sec for conf.1 is considerably larger than the value of 0.1 deg/sec for conf.2. However, this value is based on the hypothesis that the pilot's attention is optimally divided among the visual cues. Postulating for conf.1 the same threshold value for the approach angle rate information as found for the other vertical task configurations, the model analysis indicates that the subjects concentrated too much on the runway sides at the expense of the useful approach angle rate information provided by the runway threshold line.

[M] significance of differences was tested by means of a two-tailed t-test.

Finally, both observation noise spectra display an increase in observation noise for lower frequencies* which refers to the very poor approach angle information contained in both display configurations. This concrete result is in agreement with the results of previous studies (Refs. 7 and 8).

The effect of aircraft reference (i.e. pitch) information on approach performance is derived from confs. 2 and 3 (the latter displaying an aircraft reference line). Fig. 7a shows that there is no difference in approach angle (α) performance which is "explained" by the model: the aircraft reference does not provide extra information, i.e., pitch information amounts to approach angle information of the same quality (in terms of thresholds) as already available from the runway cues. The only effect is a small, but statistically significant reduction in pitch angle and control activity (smoother control). No significant frequency domain differences were found between these configurations and are therefore not shown here.

The effect of the runway centerline on the lateral approach performance can be appreciated by comparing the results of conf.4 (only the centerline) and 5 (only the runway contour). Fig.7b shows that a significantly superior lateral approach performance (in terms of the inclination angle of the runway centerline ω_c) is obtained when the centerline is explicitly visible. The explanation for this can be related to one single model parameter: the perceptual threshold of the centerline inclination angle which is substantially smaller when this centerline is explicitly displayed (table 6). The observation noise spectra given in figure 9 confirm this result. For conf.5 larger observation noise in the low-frequency region (associated with position information) is obtained.

The aircraft reference providing roll information (conf. 6) yields a significant improvement in lateral performance. This is shown in Fig. 7b and explained by the model analysis: the quality of the roll information is such that additional useful inner loop information is provided. This is also reflected by the significantly reduced mean-squared roll angle and control activity. The frequency domain measures are not given. Only the pilot describing function is affected by the additional aircraft reference corresponding with the improved performance: an increase in gain and phase lead.

Confs. 8, 9 and 10 were included in the analysis to investigate the interference between tasks utilizing the visual scene (i.e., pictorial) information. Conf.10 was included to investigate the effect of additional ground texture (tabel 1). The matching results (Ref. 1 and tables 3 and 4) support the hypothesis that there is no performance interference between the vertical and lateral visual approach task. The experimental and theoretical scores are shown in figure 2. Also the frequency domain measures are not different from the corresponding single-axis tasks and are not given here (see ref.1).

---

* In contrast with the typical flat spectrum in the lower frequency region for more ideal display situations.

Finally the perceptual thresholds resulting from the model matching procedure are considered. The threshold values are summarized in table 6 containing also the initial threshold estimates used for the model predictions. Several position threshold values are surprizingly close to the initial estimates. Furthermore, the thresholds associated with the direct perception of movement are clearly not numerically equal to the corresponding position thresholds as was initially supposed (Ref.1). Of course, the ultimate thresholds values are based on the hypothesis that the pilot's attention is optimally divided among the visual cues.
It is expected that the perceptual threshold values obtained in this study provide a useful guideline for modelling the perception process of new visual scene situations.

CONCLUDING REMARKS

A model of the visual scene perception process on the basis of linear perspective geometry and the relative (visual) motion cues was combined with the optimal control model. The model was used to predict and analyse the results of an experiment involving a variety of visual approach tasks.

Although the experimental results in terms of mean-squared performance scores agreed relatively well with the model predictions, a more detailed comparison of the frequency domain measures exhibited large differences primarily in the observation noise spectra. Therefore, an automated model matching procedure was followed.
Herewith, a matching criterion was minimized involving a weighed average discrepancy in mean-squared scores, pilot describing function data and observation noise spectral density data. Only the observation noise covariances were considered as dependent model parameters and optimized via a first-order gradient method.
By means of the efficient gradient algorithm the observation noise covariances could uniquely be identified. Under given assumptions, these covariances led to perceptual thresholds associated with the visual cues. It is expected that these perceptual thresholds provide a useful guideline for modelling the perception process of new visual scene situations.

The analysis showed that the visual scene perception model offers a general framework to deal quantitatively with the effect of (e.g.) visibility conditions, runway characteristics and aircraft reference information. In addition, the model will be useful to investigate the interaction of these characteristics with other display information, aircraft characteristics, disturbance environment, etc.

REFERENCES

1. Wewerinke, P.H.: Visual scene perception process involved in the manual approach. NLR TR 78130 U, October 1978.

2. Wewerinke, P.H.: A theoretical and experimental analysis of the outside world perception process. Paper presented at the 14th Annual Conference on Manual Control, Los Angeles, April 1978. (also NLR MP 78020 U).

3. Bryson, A.E. and Ho, Y.C.: Applied optimal control. Blaisdell Publishing Company, 1969.

4. Wewerinke, P.H.: Performance and workload analysis of in-flight helicopter missions. NLR MP 77013 U, June 1977.

5. Baron, S. and Levison, W.H.: Display analysis with the optimal control model of the human operator. Human Factors, 1977, 19(5).

6. Kleinman, D.L.: Solving the optimal attention allocation problem in manual control. IEEE Trans. on autom.control Vol.AC-21, No.6, December 1976.

7. Naish, J.M.: Control information in visual flight. Seventh Annual conference on Manual control. NASA SP-281, June 1971.

8. Havron, M.D.: Information available from natural cues during final approach and landing. Human Sciences Research Inc., HSR-RR-62/3-Mk-X, March 1962.

Table 1 Viewing conditions and selected configurations

VIEWING CONDITION	CONTROL		
	VERTICAL	LATERAL	BOTH
	1		
		4	
	2	5	8
	3	6	9
			10

N.B. CONFIGURATION 7 CONCERNS ROLL TRACKING BASED ON ROLL BAR ONLY

Table 2 Weightings in the matching criterion

weighting	value	total
a_{α,ω_c}	0.6	1.05
$a_{\theta,\varphi}$	0.3	
a_{δ_e,δ_a}	0.15	
a_{yp}	0.5	1.0
a_{on}	0.5	

Table 3 Matching results - vertical control tasks

a) The best model match

CONF.	MATCHING CRITERION				PERCEPTUAL THRESHOLD				
	L_{ms}	L_{yp}	L_{on}	L_t	a_{ω_1}	$a_{\dot{\alpha}_1}$	a_{α}	$a_{\dot{\alpha}}$	$a_{\tilde{\theta}}$
1	0.03	0.30	0.34	0.67	5.2	0.23			
2	0.10	0.45	0.25	0.80			0.51	0.11	
3	0.04	0.50	0.30	0.84				0.10	0.75
8V	0.11	0.43	0.31	0.85			0.49	0.14	
9V	0.10	0.68	0.25	1.03			0.44	0.08	0.64
10V	0.24	0.52	0.20	0.96			0.48	0.02	0.69
Average	0.10	0.48	0.28	0.86	5.2	0.23	0.5	0.1	0.7

b) The ultimate model match

CONF.	MATCHING CRITERION			
	L_{ms}	L_{yp}	L_{on}	L_t
1	0.03	0.30	0.34	0.67
2	0.10	0.45	0.25	0.80
3	0.12	0.50	0.30	0.92
8V	0.58	0.47	0.26	1.31
9V	0.66	0.66	0.24	1.56
10V	0.47	0.52	0.17	1.16
Average	0.33	0.48	0.26	1.07

THRES-HOLD	AVERAGE VALUE
a_{ω_1}	5.2
$a_{\dot{\alpha}_1}$	0.23
a_{α}	0.5
$a_{\dot{\alpha}}$	0.1
$a_{\tilde{\theta}}$	0.7

Table 4 Matching results - lateral control tasks

a) The best model match

CONF.	MATCHING CRITERION				PERCEPTUAL THRESHOLD					
	L_{ms}	L_{yp}	L_{on}	L_t	$a_{\omega_{c_4}}$	$a_{\dot{\omega}_{c_4}}$	a_{ω_c}	$a_{\dot{\omega}_c}$	a_{φ}	$a_{\dot{\varphi}}$
4	0.08	0.46	0.35	0.89	0.4	2.5				
5	0.09	0.43	0.39	0.91			1.5	2.2		
6	0.18	0.58	0.33	1.08			2.4		1.0	
7	0.32	0.65	1.07	2.04					1.1	3.6
8L	0.61	0.56	0.53	1.70			1.8	3.1		
9L	0.18	0.37	0.39	0.94			1.7		1.7	
10L	0.15	0.55	0.43	1.13			1.8		1.7	
Average	0.23	0.51	0.50	1.24	0.5	2.5	1.8	2.5	1.2	3.5

b) The ultimate model match

CONF.	MATCHING CRITERION			
	L_{ms}	L_{yp}	L_{on}	L_t
4	0.13	0.46	0.37	0.96
5	0.82	0.51	0.45	1.78
6	0.50	0.57	0.36	1.43
7	0.46	0.70	1.12	2.28
8L	1.14	0.43	0.48	2.05
9L	0.53	0.33	0.34	1.20
10L	0.96	0.46	0.32	1.71
Average	0.65	0.49	0.49	1.63

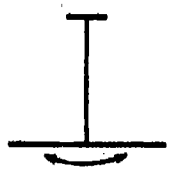
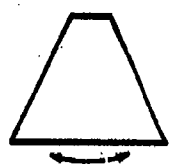
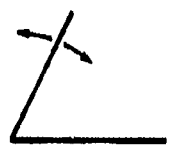
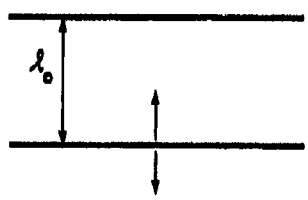
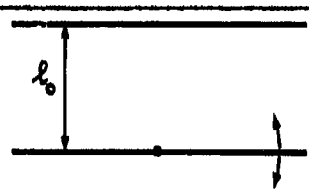
THRESHOLD	AVERAGE VALUE
$a_{\omega_{c_4}}$	0.5
$a_{\dot{\omega}_{c_4}}$	2.5
a_{ω_c}	1.8
$a_{\dot{\omega}_c}$	2.5
a_{φ}	1.2
$a_{\dot{\varphi}}$	3.5

Table 5 Matched model and measured scores

vertical configuration		$MS_{\alpha}(deg^2)$	$MS_{\theta}(deg^2)$	$MS_{\delta_e}(N^2)$
1	model	0.194	0.098	27.0
	measured	0.193	0.098	26.4
2	model	0.076	0.050	22.2
	measured	0.077	0.054	23.0
3	model	0.072	0.046	21.8
	measured	0.081	0.047	21.2
8v	model	0.074	0.050	22.2
	measured	0.083	0.061	24.6
9v	model	0.073	0.048	22.0
	measured	0.058	0.042	19.8
10v	model	0.073	0.048	22.0
	measured	0.072	0.038	20.3

lateral and roll configuration		$MS_{\omega_a}(deg^2)$	$MS_{\phi}(deg^2)$	$MS_{\delta_a}(N^2)$
4	model	3.76	7.86	10.5
	measured	3.62	7.82	10.9
5	model	5.70	9.05	10.9
	measured	4.72	7.90	10.9
6	model	3.40	5.81	9.74
	measured	3.99	5.40	9.43
7	model	-	3.10	4.29
	measured	-	2.99	4.99
8L	model	5.21	8.68	10.7
	measured	6.20	12.1	14.0
9L	model	3.40	5.82	9.75
	measured	4.55	6.73	10.5
10L	model	3.40	5.81	9.74
	measured	4.49	7.07	11.1

Table 6 Perceptual thresholds from the model match

Par	display	threshold "matched"	initial threshold estimate
		0.5° 2.5 °/sec	1° 1 °/sec
ω_c $\dot{\omega}_c$		1.8° 2.5 °/sec	2° 2 °/sec
		5.2° -	5° 3 °/sec
θ $\dot{\theta}$		0.7° -	0.5° 0.3 °/sec
α $\dot{\alpha}$		0.5° 0.1 °/sec (0.23 °/sec for conf.1)	0.5° 0.3 °/sec
ϕ $\dot{\phi}$		1.2° 3.5 °/sec	0.7° 1 °/sec

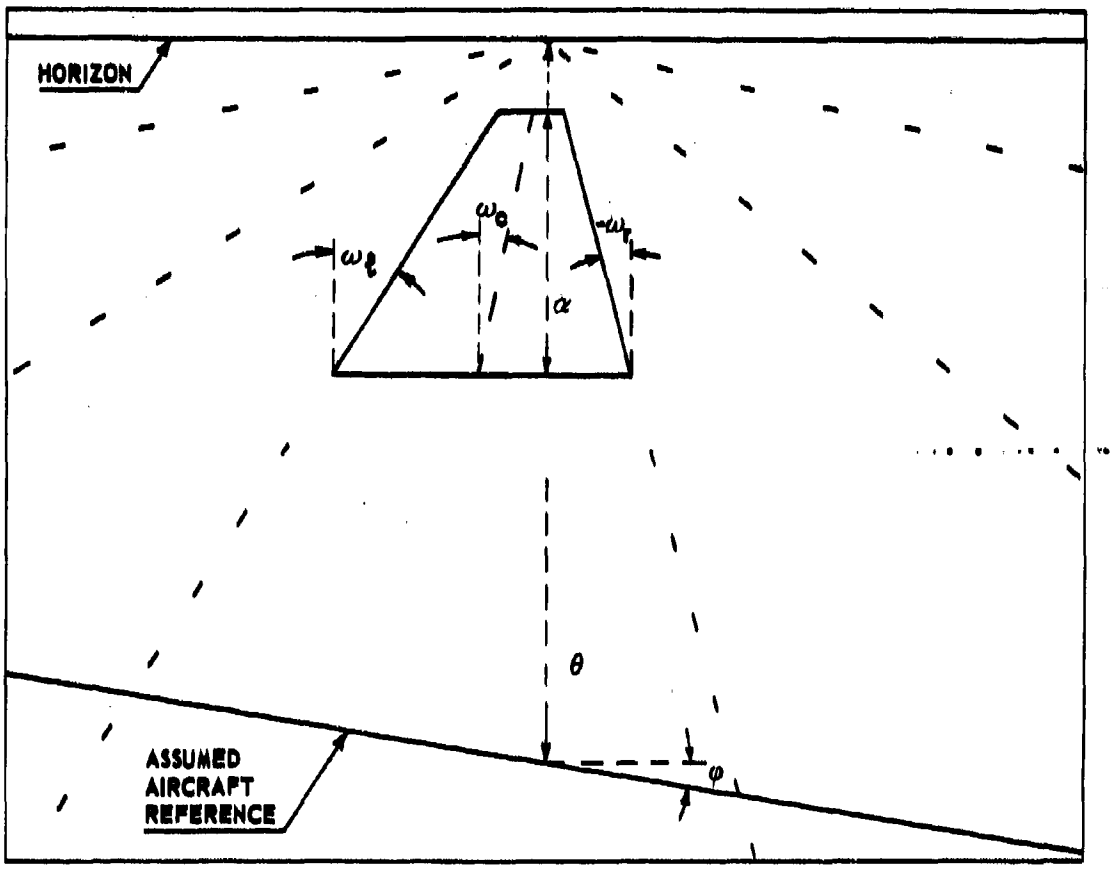


Fig. 1 Cues derived from the visual approach scene

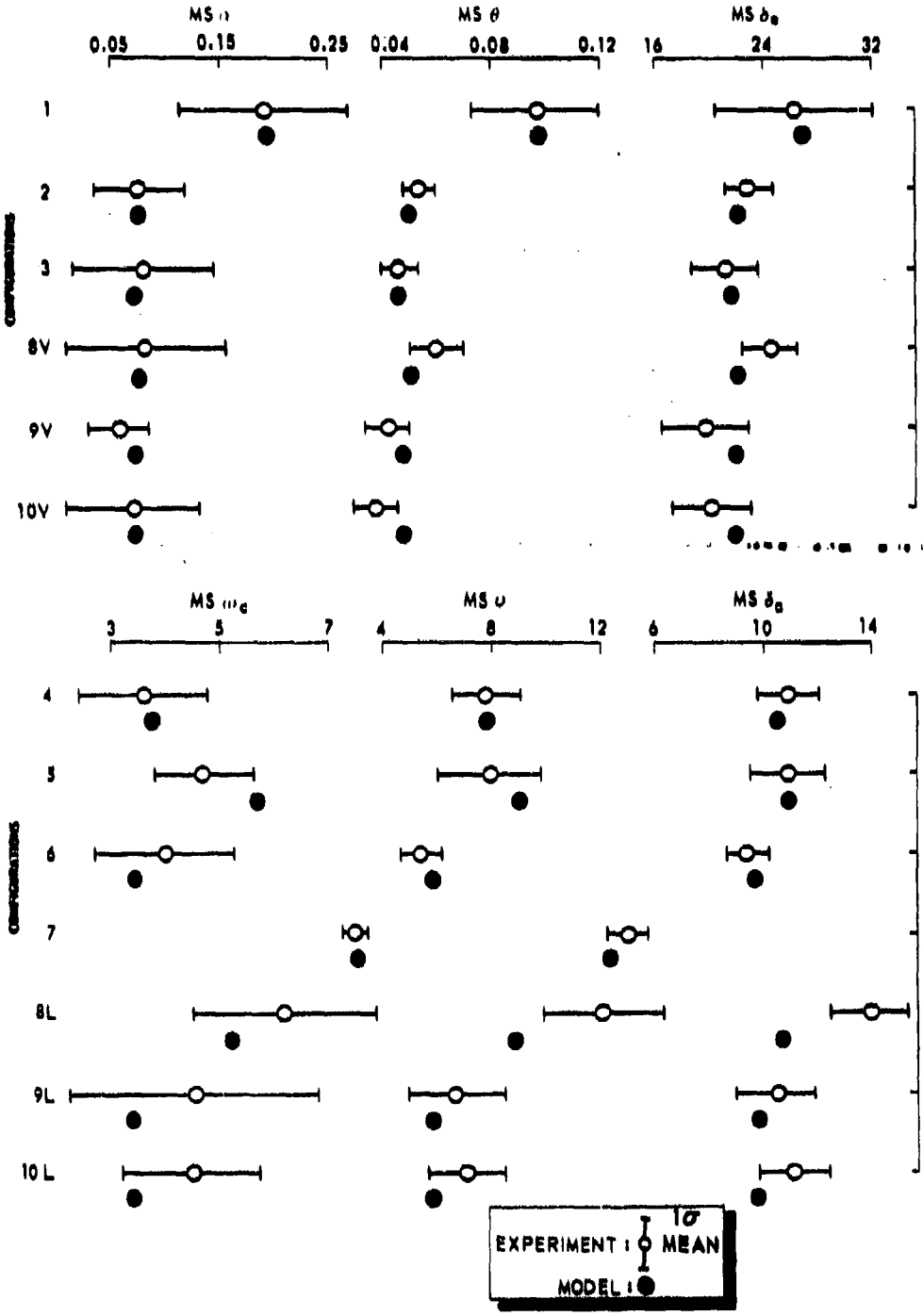


Fig. 2 Comparison of experimental scores and model match

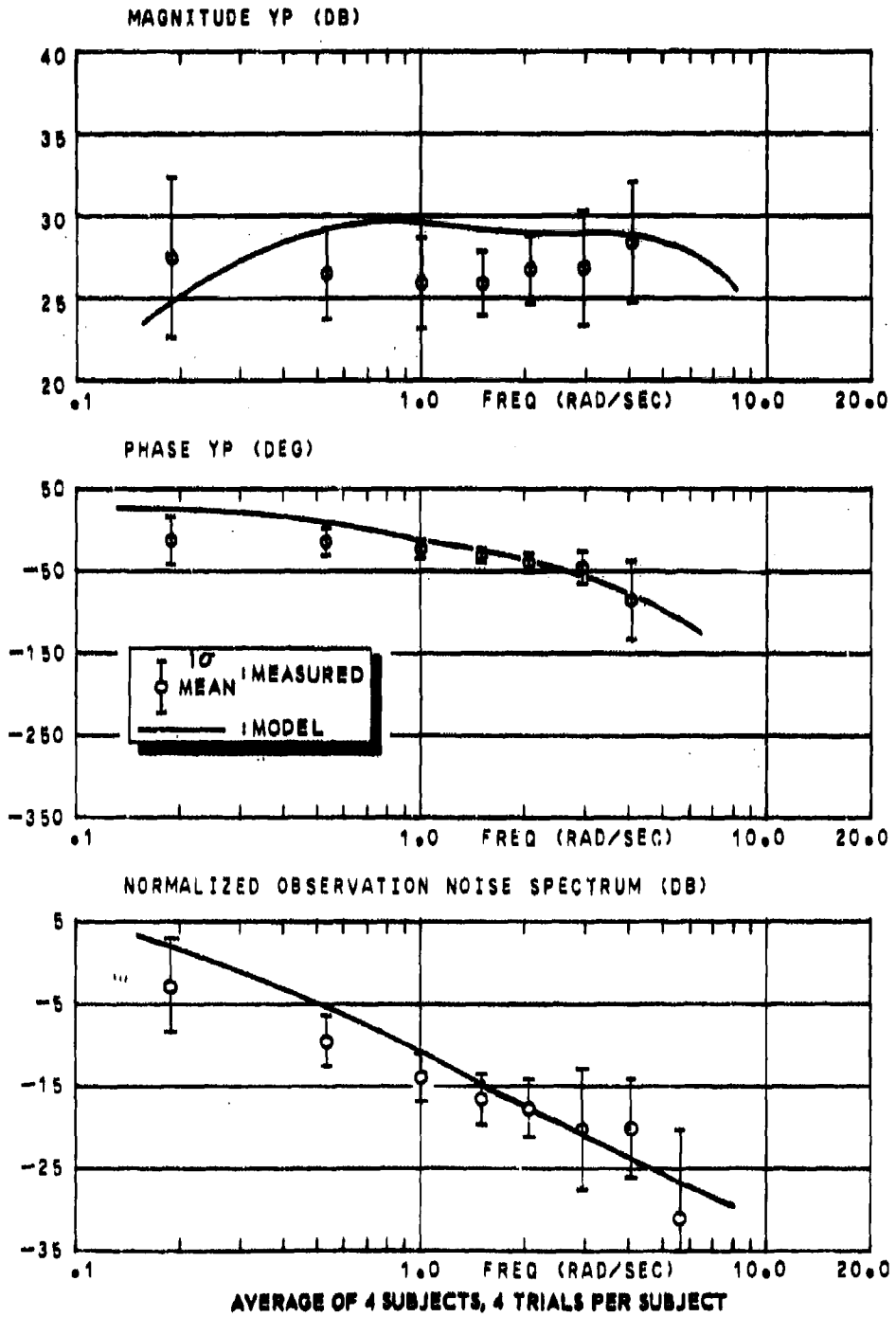


Fig. 3 A comparison of measured and model frequency domain measures - configuration 1

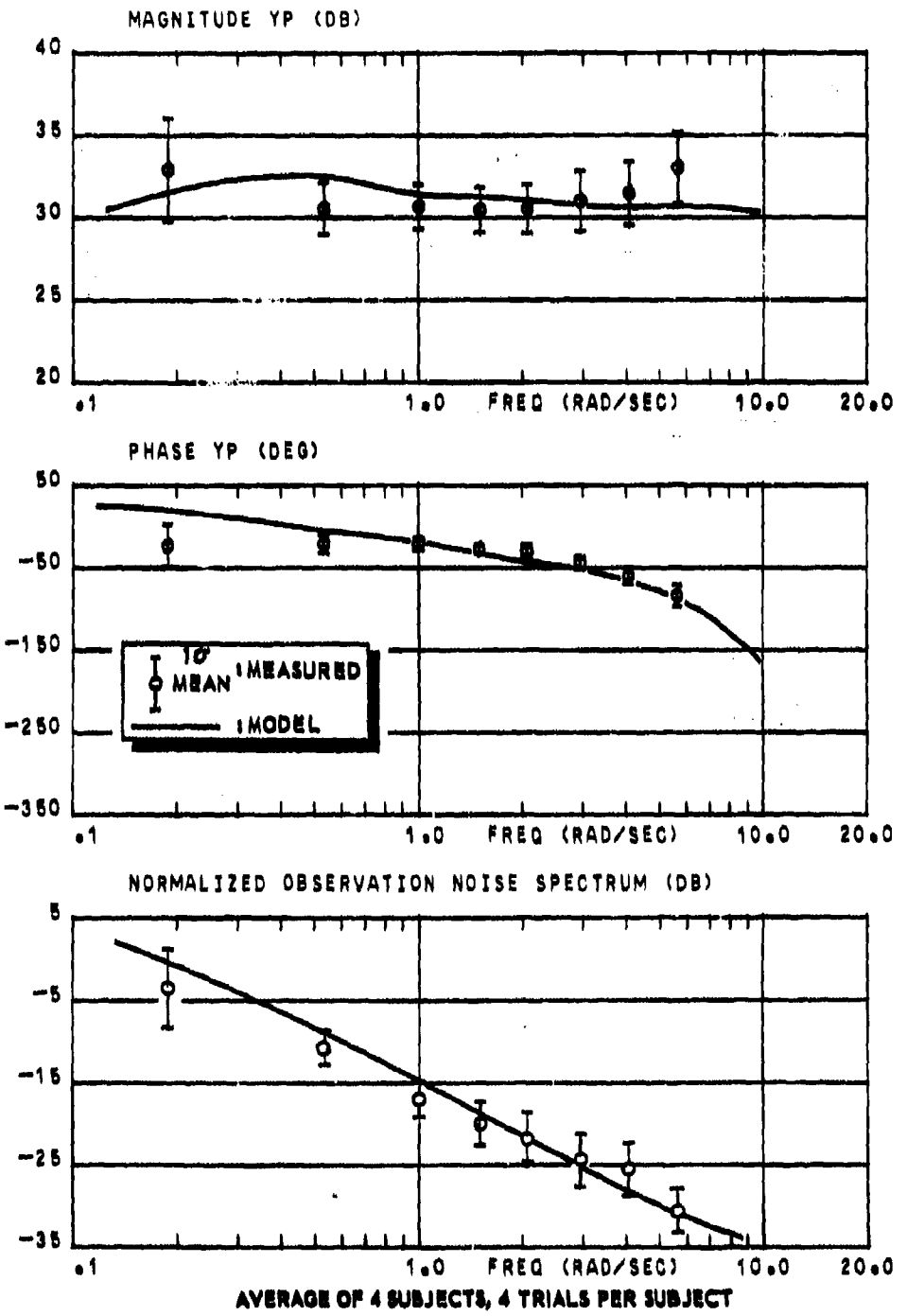


Fig. 4 A comparison of measured and model frequency domain measures - configuration 2

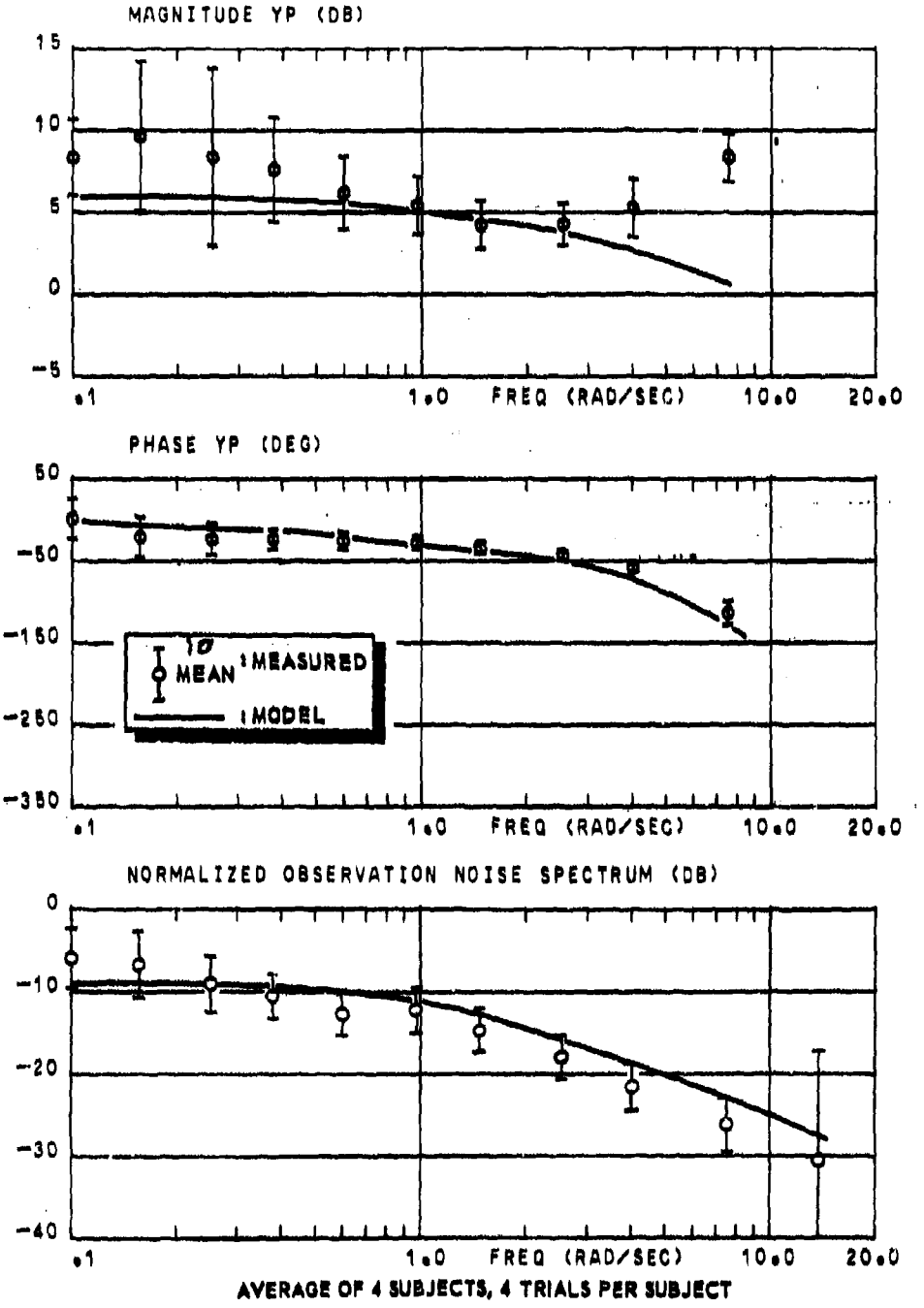


Fig. 5 A comparison of measured and model frequency domain measures - configuration 4

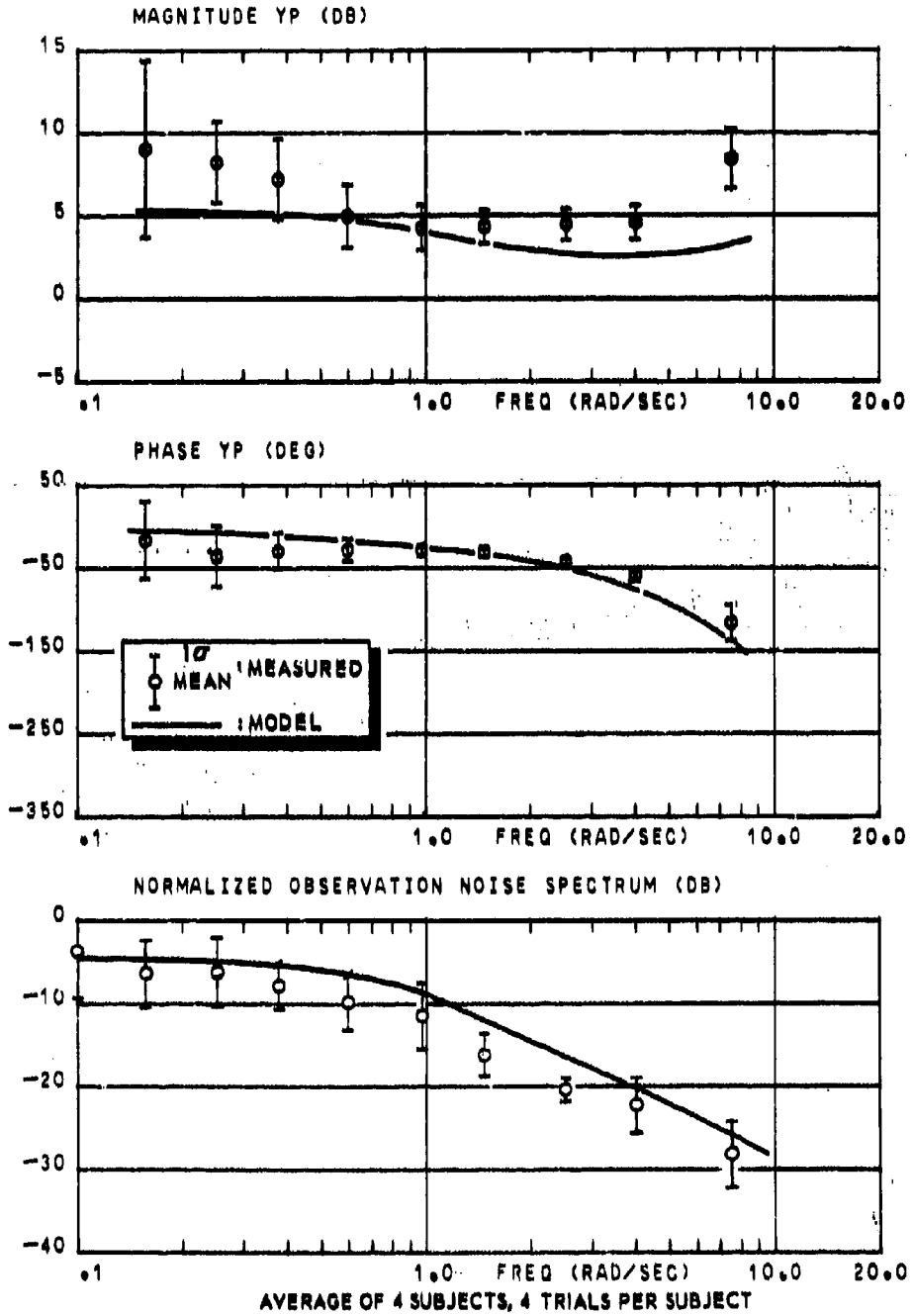


Fig. 6 A comparison of measured and model frequency domain measures - configuration 5

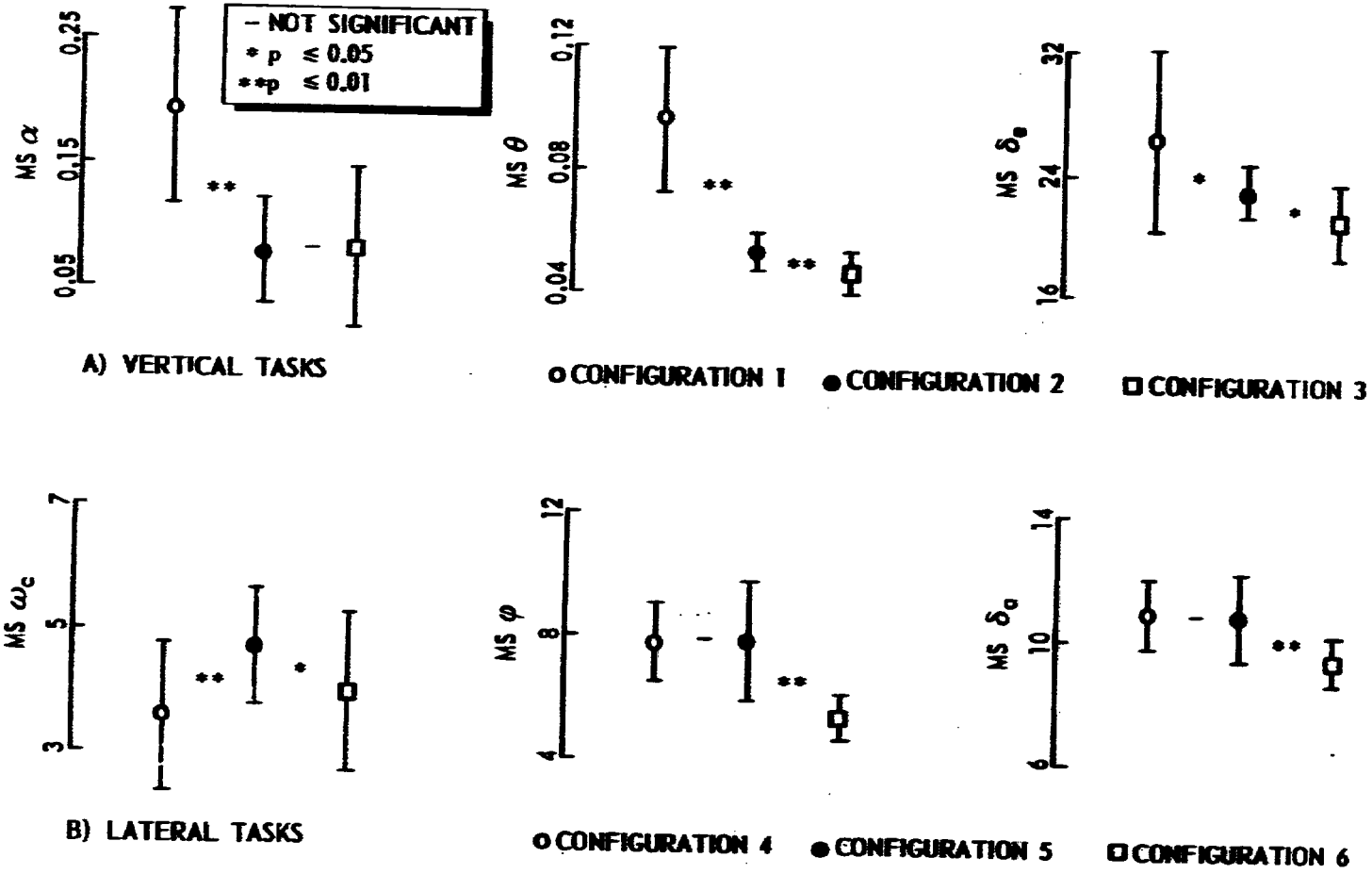


Fig. 7 The effect of visual scene characteristics on approach performance

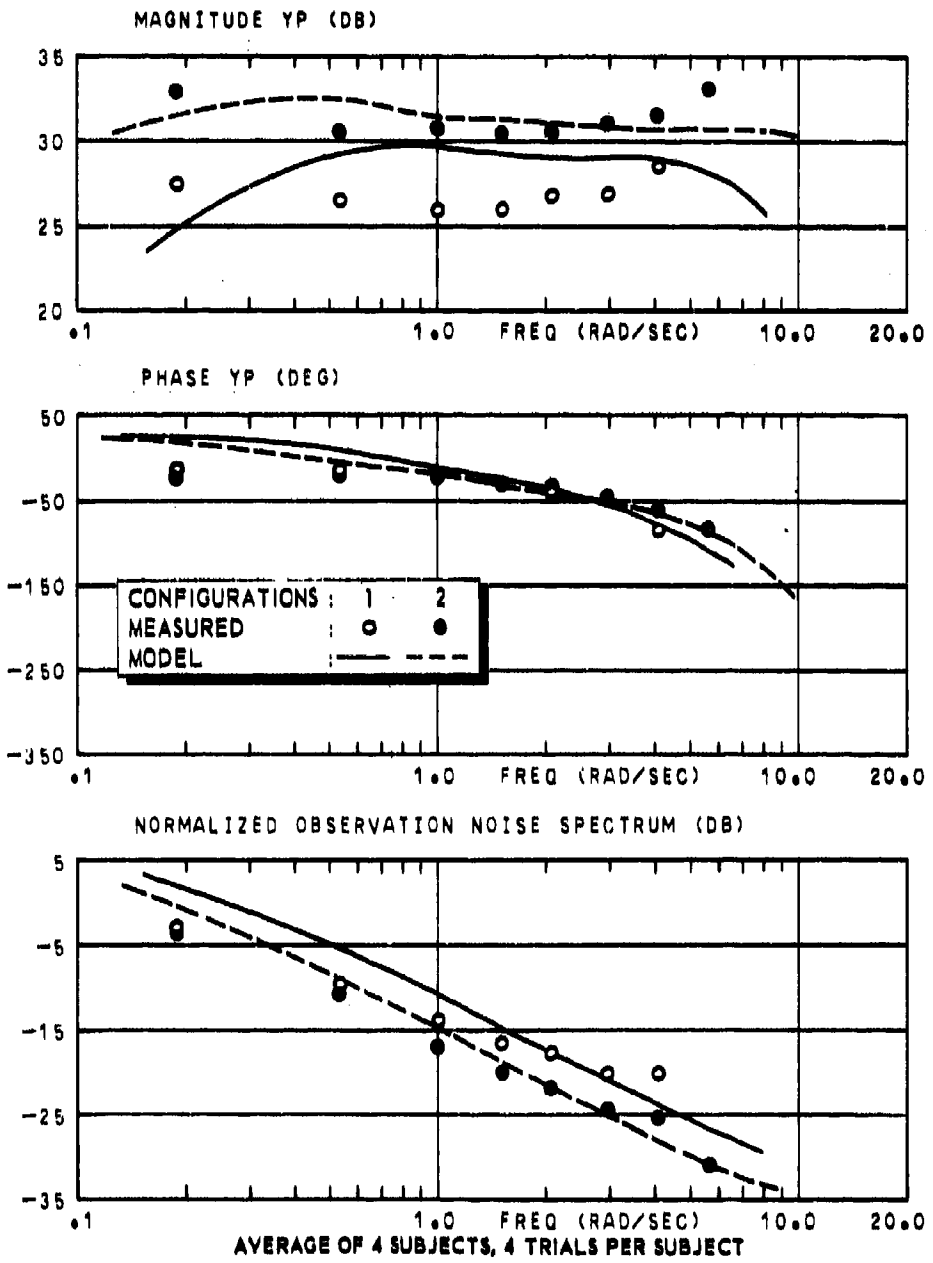


Fig. 8 Frequency domain results for confs. 1 and 2 (runway visual range-effect)

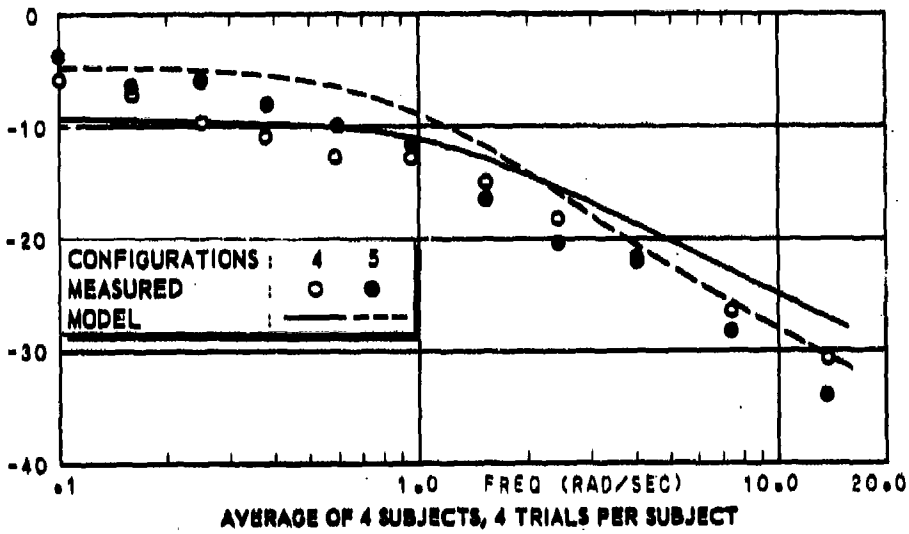


Fig. 9 Normalized observation noise spectra for confs. 4 and 5 (runway centerline-effect)

A STUDY OF THE USE OF COLOR GRAPHICS FOR REPRESENTING SCHEMATIC INFORMATION*

Heidi Neubauer
Department of Computer Science

William B. Rouse
Department of Mechanical and Industrial Engineering
Coordinated Science Laboratory
University of Illinois, Urbana, Illinois, 61801

ABSTRACT

It is proposed that graphical computer displays can aid human operators of complex systems. Further it is proposed that as systems become complex, the use of color can help. ...In this informal paper the results of the first two experiments in this area are presented. Here abstract schematic networks are used to study responses about the relationships of nodes. Two sizes of networks and various colorings are used. In the first experiment it is asked whether a relationship between two nodes exists. In the second experiment it is asked which of another set of nodes holds a functional relationship to the others.

INTRODUCTION

The research reported in this informal paper is aimed towards an eventual representation of procedural information on graphical computer displays. We started by looking at emergency and abnormal procedure descriptions for the DC-10 aircraft since we had a complete set of on-board manuals available. What is striking is a potential problem in the desired rapid and efficient use of this type of manual. The systems (fuel, hydraulic, electrical, etc.) are depicted schematically in 3 or 4 color fold-out charts, while the procedures to be followed concerning these systems are outlined in often pages-long verbal descriptions of essentially algorithmic procedures. Aircraft are not the only area where this duality occurs. Process control plants and nuclear power plants are other examples where operations and procedures about spatially oriented systems of objects are described in written form.

*This work was supported by the National Aeronautics and Space Administration under NASA-Ames Grant NSG-2119.

We see two problems with this technique of describing such procedures purely textually, or as in the example above, textually with the schematic in a separate section of the manual or, in fact, in another manual. First, once the problem is detected two transformations are needed to complete the required task. The operator makes an observation about the physical spatially-oriented system (in any of many ways), transforms his observation into the appropriate reference in the textual manual, and then transforms once again to perform the spatially-oriented action required by the physical system. Second, a text description does not give the feeling for a spatial procedure that a well-constructed graphical display might. It is clear that the verbal text can offer greater precision in fine description though at the expense of large amounts of time to read and absorb minutiae. We feel that it may prove fruitful to investigate the performance in terms of speed and accuracy which can be obtained with graphical representations of systems. Additionally, as these systems become complex, we feel that the relationships among components can become difficult to perceive and that the use of color can help understanding.

The display used in these experiments is an abstraction of various systems -- a square grid of 9 or 25 interconnected nodes (subsystems) with inputs to the left and outputs to the right. Thus, by outputs from some nodes becoming inputs to others a system can be represented which is not unlike the diagrams seen in the aircraft manuals mentioned above. See Figure 1 for an example of the sort of display we used, reproduced here without colors.

In the first set of experiments the task subjects performed was to answer whether or not the node labelled 'A' affected the node labelled 'B'. The 'A' node was always placed in the leftmost column, column 1, while the 'B' node was always in column 5. None of the other nodes were labelled, and no other information besides the schematic and the two letters appeared on the display. Four different coloring schemes were used with the two sizes of networks mentioned above. The second set of experiments asked subjects to determine which of two intermediate nodes (valves or switches) would disconnect 'A' from 'B'.

EXPERIMENT 1

Two factors were investigated -- size (and therefore complexity) of the network and color coding scheme. Networks were always displayed against a black background. The first of the four color coding schemes was monochrome white to mimic the standard black and white, or green and white graphic cathode-ray tube display. The second color coding was a two color scheme of white and red where all of the display was white except for those output lines which were 'feedback' paths. Feedforward was defined as paths progressing left to right and up to down across

the network. An earlier set of experiments by Rouse showed that subjects had difficulty with the feedback paths [3]. We thought that if these paths were accentuated by a red line that they might not cause confusion. The third color scheme was a n-color scheme where n was the number of nodes in the network. Each node had its own unique and distinct color, and every output from that node was that color. The set of colors used in each display comes from what is known as the "palette". The choice of colors will be described in greater detail in a later section of this paper. The fourth was a randomized scheme, where each line segment was drawn with a color randomly selected from the appropriate color palette for the network size. In these displays, the consistency of coloring over the path was lost, since each path was drawn of up to 5 line segments. However, the colors did help to discriminate line crossings. Thus, this coding supposedly de-emphasized relationships among components while maintaining discriminability.

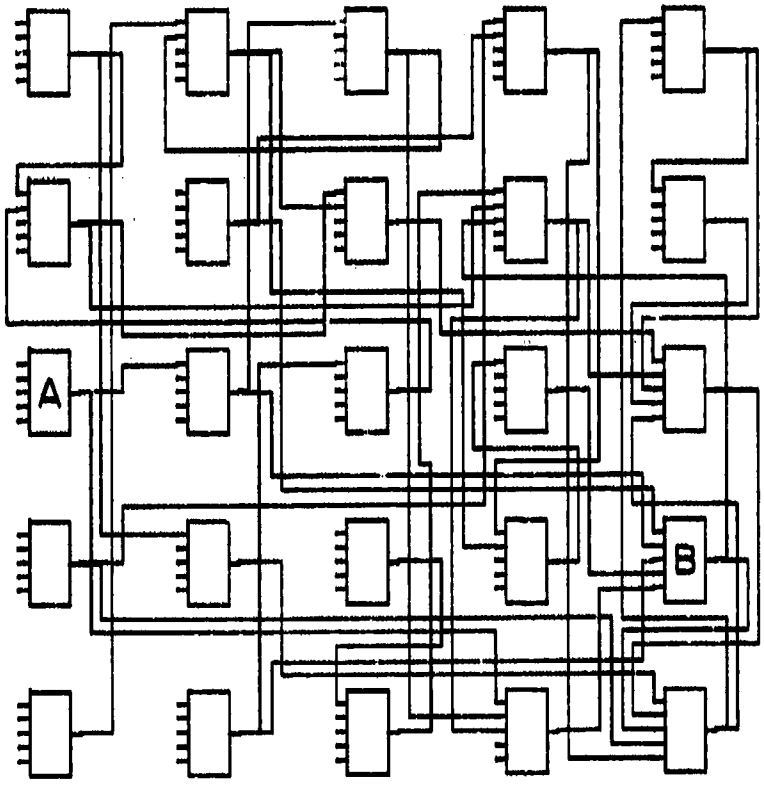


Figure 1.
Example Display from Experiment 1
"Does A Affect B?"

Ten unpaid engineering graduate and undergraduate students served as subjects, nine males and 1 female. Each saw 80 displays -- 10 of the 5 x 5 and 10 of the 3 x 3 in each of the four colorings outlined above. The order of display presentation was randomized once, and then presented to each subject in the same random order. A joystick served as the input device. A deflection to the right was recorded as a yes and to the left, a no. Each session lasted about an hour with perhaps half the time taken in the generation of displays. The response time, in seconds, and the correctness of the answer were recorded and analyzed.

The results of this first experiment are shown in the table below.

TABLE 1
EXPERIMENT 1, MEAN RESPONSE TIME (seconds)

	white	colors	feedbk	rand
3x3	9.11	7.61	10.62	10.30
5x5	12.02	10.29	13.04	15.24

The analysis of variance of mean response time showed that size was significant ($p < 0.001$, $F_{1,72} = 12.66$) as was color coding ($p < 0.05$, $F_{3,72} = 3.30$). Interaction between the two factors was not significant. Analysis of the percentage of correct responses showed no significance. Since the subjects were instructed to "work quickly but accurately, and to not guess," we assume that they worked to a criterion of perfection, taking as much time as they needed to be certain. The mean scores for percentage correct ranged from a low of 94 to a high of 100 with an overall average of 96.8. The data was also analysed with an additional factor of time. There were 2 levels of this factor, the first five of the displays seen of a given coloring and size, and the last five. The results of this analysis of variance show that time was significant ($p < 0.001$, $F_{1,64} = 15.537$). We assume that subjects learned to perform faster, maintaining high accuracy as the experiment progressed.

EXPERIMENT 2

This experiment was concerned with relationships between internal network nodes and 'A' and 'B'. A scenario was built into the directions:

"The node labelled 'A' always connects to the node labelled 'B'. Imagine 'A' is a fuel tank and 'B' is an engine on fire. Imagine the lines carry fuel. Two valves labelled 'L' (left) and 'R' (right) are available but only one will disconnect 'A' from 'B' if activated."

The nodes labelled 'A' and 'B' appeared in the same columns as experiment 1, while the new labels, 'L' and 'R' always appeared in columns 2 and 4, respectively. See Figure 2 for an example of this display. The experimental procedure followed was similar to the first. Subjects determined which valve would disconnect 'A' from 'B' and then move the joystick to the right for the right valve, and to the left for the left valve. However, the coloring schemes differed, and the order of color scheme presentation was blocked in a Latin square fashion to eliminate ordering and condition effects. Twelve unpaid subjects saw 4 sets of 10 displays, each set with a coloring determined by the Latin square design. There were then 4 orderings, and 4 sessions of the same 10 displays (though each session was ordered differently). With twelve subjects, there were three replicates of the complete Latin square design. The color palettes used had 1, 2, 4, or 8 colors -- each smaller set of colors being a proper subset of the next. The response time, in seconds and the correctness of the answer were recorded but only the response time was analyzed in view of the results of experiment 1.

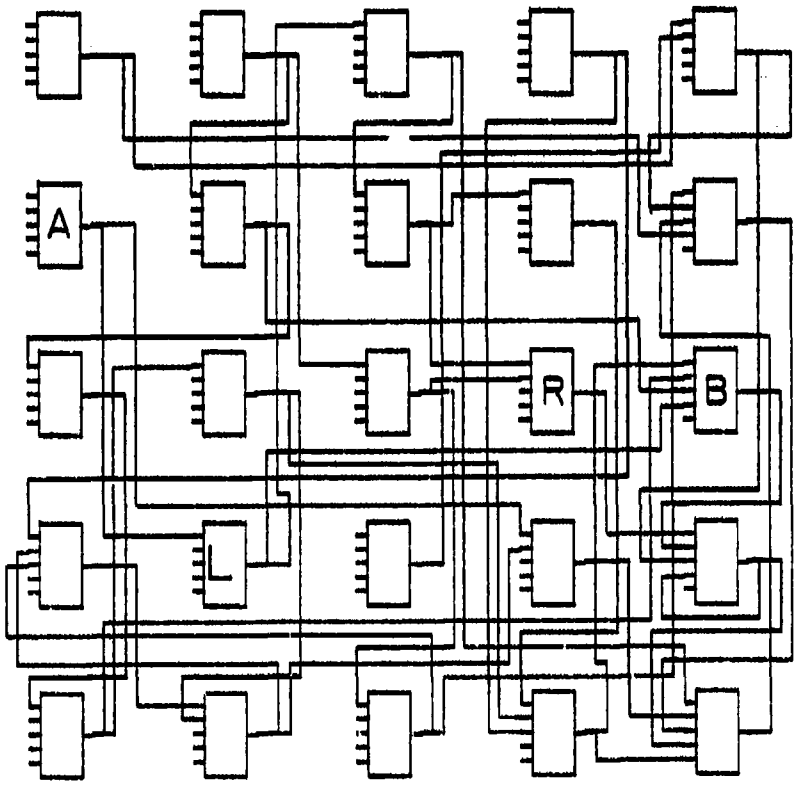


Figure 2.
Example Display from Experiment 2
"Which Valve, L or R, Will Shut Off the Flow from A to B"

TABLE 2
EXPERIMENT 2, COLORINGS AND TREATMENT MEANS

Palette Name	Treatment Mean	Color(s) in Displays
w	9.95	white
2	8.65	white, red
4	8.78	white, red, blue, green
8	8.09	white, red, blue, green, grey, pink, orange, yellow

TABLE 3
EXPERIMENT 2, LATIN SQUARE DESIGN
MEAN RESPONSE TIME

				Sessions				Row
Palette		Order		1	2	3	4	Means
w	8	2	4	14.3	8.7	8.7	7.6	9.71
2	w	4	8	9.5	7.7	6.9	5.2	7.33
4	2	8	w	10.8	8.4	6.9	5.2	8.53
8	4	w	2	11.5	9.8	9.7	8.6	9.91
Column Means				11.55	8.66	7.92	7.34	

The results show that the ordering of palette groups (rows in the table above) is significant ($p < 0.05$, $F_{3,36} = 3.383$). The group of three subjects who were presented the displays in the order 2, w, 4 and 8 colorings had the shortest average time not only for the whole session but also for each of the subsessions. We might ask whether performance on a display with 8 colors is facilitated by earlier training on more displays with fewer colors. Recall that the displays in each of the cells were exactly the same except being colored and ordered (within each session) differently.

The 4 sessions (chronological sequence of sessions) also showed significance ($p < 0.001$, $F_{3,36} = 8.299$). Each cell (with one exception) shows a decrease in the mean response time from the previous cell, regardless of coloring. This confirms many of the subjects' reporting that they began to recognize the fact the networks were being repeated by noticing the patterns with which the labels 'A', 'L', 'R', and 'B' appeared. Treatments, the different color codings, did not show significance. The blocking of treatments -- color codings -- may make it easier to respond quickly. Recall that in experiment 1 where displays were presented in random order, there was a significant effect of coloring.

FUTURE PLANS

The assignment of colors to nodes and outputs has been arbitrary, simply one color per node, or a set of colors repeated sequentially. The meaningfullness of this scheme is limited. We know that colors themselves carry many

multidimensional attributes (e.g. red -- danger, hot, immediate, stop). In addition, we would like to investigate coding a given node and other chosen nodes in some fashion to emphasize their relationships. We would like to build up a set of experimental displays and scenarios so that we could ask subjects to intervene in the control of a graphically displayed system. Subjects could actually enter commands to institute a needed procedure and use the graphics and colors on the display as their prompts.

Further data analysis can be done on experiment 2. The displays were selected so that the minimal path from node 'A' to node 'B' was known to pass through 3 to 5 nodes. Of the 10 different displays shown, 4 had paths of length 3 another 4 of length 5, and 2 of length 2. Half of each set had correct answers of 'left' and half of 'right'. We want to compare the time to answer questions about short pathed (3 nodes) versus long pathed (5 nodes) display relationships.

APPENDIX - COLORS

The colors used in the 3 x 3 networks in experiment 1 and the 1,2,4 and 3 colored 5 x 5 networks of experiment 2 were taken from the set of eleven basic color terms described by Marvis in [1]. Given the Munsell notation of the desired colors, we obtained the actual glossy finished color chip [2], and attempted to replicate the color on a Ramtek 9400 Color Display System. The Ramtek allows the 3 color guns to be adjusted by software control. The small paper chip was held up to the screen, with the display's background held a relatively constant grey, and the palette entry adjusted until the match between the chip and the screen was subjectively best. The difference in the media (luminous CRT screen and opaque color chip) is great, and a problem in describing the effective stimuli used in this type of research. The 25 color palette was chosen arbitrarily, but with consideration to the even placement of colors on the color wheel. The eleven Munsell colors did not seem appropriate for inclusion in this larger palette because their discriminability was too great. When other colors were added to the Munsell palette to enlarge it, the others seem to fade while the Munsell set stood out. Therefore we chose a different set of 25 colors so that all were discriminable but none was distracting.

REFERENCES

1. Mervis, Carolyn B., Jack Catlin and Eleanor Rosch, "Development of the Structure of Color Categories," Developmental Psychology, Vol 11, No. 1, 1975, pp. 54-60.

2. Munsell Color Book, Munsell Color Company, Baltimore, Maryland.

3. Rouse, William B., "A Model of Human Decisionmaking in Fault Diagnosis Tasks that Include Feedback and Redundancy," IEEE-Systems, Man and Cybernetics, Vol SMC-9, Number 4, April, 1979,pp. 237-241.

CONTENT, SYMBOLOGY, AND FORMAT OF COCKPIT DISPLAY OF TRAFFIC INFORMATION: GENERAL AVIATION PILOT OPINION

Sandra G. Hart *
Tufts University
Medford, Mass. 02155

ABSTRACT

In order to evaluate the concept of cockpit display of traffic information (CDTI) and its impact on the air traffic control system, a reasonable candidate display will be required for research. This study was designed to provide guidelines for later simulation research on display design. Because pilots will be the ultimate users of CDTI, 13 general aviation pilots were asked to respond to 250 questions about display content, format and symbology. The 100 static displays shown as examples incorporated information that might be included in a CDTI (e. g. navigation, terrain, weather, and status and position of other aircraft) presented with varying complexity, symbology, and format. The general aviation pilots' responses were remarkably similar to those previously obtained from airline pilots. The few differences found seemed to reflect the general aviation pilots' concern with economics and a single pilot's workload under IMC. They felt that a lot of information should be available with pilot control over the amount of information displayed at any one time.

INTRODUCTION

In the near future, technological advances in automation, display technology, and information dissemination will make it possible to provide computer-generated cockpit displays of traffic information (CDTI) via data link. A CDTI should provide a pilot with an awareness of his own situation and that of other nearby aircraft which would allow him to detect loss of separation and ATC or airborne system failures. In addition, a CDTI could assist the pilot in evaluating the intentions of other aircraft and allow him to anticipate and plan ahead. It could also relate the position of a pilot's own aircraft and that of others to the environment. The presentation of such information in the cockpit could result in substantial changes in the role of the pilot in the air traffic control system. It has been suggested that certain air traffic controller functions such as intrail spacing, terminal area merging, and acting as an ATC system backup and monitor could be assumed by pilots. This concept of redefining the roles of pilot and controller has been referred to as distributed management.

To perform any type of ATC functions from the cockpit, it may be necessary to display information about the environment (e.g. weather, terrain, and navigation routes) in addition to the status, intent, and position of other aircraft. Pilots now obtain such information from a variety of sources including ATC, charts, weather radar, monitoring the radio transmissions of

* This research was conducted at NASA-Ames Research center and was sponsored by NASA grant NSG-2156 to Tufts University.

other aircraft, and looking out of the window. A CDTI could integrate these different types of information to assist pilots in forming a mental picture of the environment. Many questions remain, however, concerning what information should be displayed, in what format it should be presented, and with what level of complexity.

Although previous research into the concept of providing pilots with traffic information has shown that the information provided enhances control in longitudinal spacing and awareness of the traffic situation (refs. 1, 2, and 3), many questions remain to be answered about the impact of CDTI on pilot and controller workload, the functions that a pilot might perform with a CDTI, and the pilot's ability to assess such a display while performing his primary duties.

In order to demonstrate the feasibility of the concept of CDTI, extensive laboratory and simulation research must be conducted. A prerequisite of the research is a candidate cockpit display. The design of the display(s) used for research purposes could have considerable impact on the validity of the conclusions drawn about the feasibility of the CDTI concept itself. Because pilots will be the ultimate users of CDTI, it was felt that they should be involved in the initial stages of designing the displays to be used in research. To this end, a number of candidate displays were simulated with a computer graphics system that incorporated different categories of information presented with varying levels of complexity, symbologies, and formats. Initially, the candidate displays and display features were shown to 23 airline pilots who were asked to evaluate the specific features that they were shown individually and in combination with other display elements (ref. 4). These pilots were asked to specify the display content and format that would incorporate all of the essential information presented in an optimal fashion with minimal display clutter and confusion.

In the current study, the same displays were shown to a group of general aviation pilots who were asked to respond to the same set of questions in order to compare the responses of the two groups of pilots. The purpose of these two studies was to obtain a set of guidelines, provided by the user-population, which would be used as the basis for subsequent laboratory and simulation research. It was not intended that this study would result in the design of a single, optimal CDTI display, but rather would define the information that a significant number of pilots felt should be incorporated into a CDTI with preferred symbology and display format. These recommendations could then be used to guide laboratory research in which the display features that pilots preferred would be tested with dynamic simulations to verify that, in fact, the preferred features did allow accurate and efficient performance with acceptable levels of pilot workload.

METHOD

The procedures, survey format, and display content have been described in detail in an earlier report (ref. 4) and thus will be presented in summary form in the present report.

Subjects

Thirteen general aviation pilots, five women and eight men, served as paid participants in the study. Their ages ranged from 30 to 60 years and their total flight hours ranged from 2,000 to 16,000 hours. All of the pilots were instrument-rated and had flown a variety of one and two engine light aircraft. Only two of them had previously participated in research at NASA-Ames Research Center, and none of them had been involved in research on CDTI. When the pilots were being sought for the study, they were not told that CDTI was involved to avoid any selection bias. Although only 15% of them were somewhat familiar with the concept of CDTI, 85% of them reported in advance of the study in a preliminary questionnaire that the addition of a traffic display in the cockpit might provide useful information.

Procedure

Upon their arrival the pilots were seated around a conference table so that each of them was able to see the projected displays. After filling in a brief summary of their professional experience, instructions were read that described CDTI, the environment into which it would be introduced, the tasks that a pilot might be asked to perform with a CDTI, and the impact that it might have on the division of responsibility between the air and ground. Each pilot was provided with a 12-page booklet in which the items that they would be asked to evaluate, directions about how to respond, and spaces for their responses and comments were organized into different sections. Several response formats were used: (1) yes or no; (2) acceptable or unacceptable; (3) necessary, desirable but not necessary, unnecessary, and undesirable; and (4) selecting a single preferred option from a set of options. Most of the items were rated individually and in comparison to other options as well.

Presentation of the displays and the pilots' responses were divided into seven segments in which display content, format and symbology for a specific category of information were investigated: (1) navigation; (2) terrain; (3) weather; (4) own aircraft; (5) other aircraft (rules for display, symbology, coding schemes, data blocks); (6) use of color coding; and (7) general questions. A brief introduction was read to the pilots and all of the displays for that category were presented before the pilots responded to the questions in that section. They were allowed to ask questions at any time and were shown the displays a second time if they requested. The pilots were given as much time to respond as they required and were encouraged to make written comments, additions, substitutions, and deletions to the display examples that they had been shown in order to develop a set of display specifications that was representative of their opinions.

Test Materials

More than 100 sample displays were created to depict different CDTI concepts individually and in the context of a basic navigation display alone or in combination with other environmental information. Display content, symbology, and format were varied for each category of information with and without color coding of individual elements. The displays were drawn by the author with a magnetic pen and pad input device to an Evans and Sutherland Picture System II. This stroke drawing calligraphic system provided five colors, of which only red, green and yellow were used for maximum contrast (ref. 5). In essence,

the displays that were created served as a series of "straw men" for which the pilots were encouraged to suggest additions, substitutions, and deletions.

The scenario upon which the design of the displays was based represented a standard southern approach to runway 30L at San Jose Municipal Airport in California. The approach was simulated for a medium jet at an initial distance of 50 miles from the airport at an altitude of 15,000 ft, heading 302° with a ground speed of 280 kts. Six different map scales were simulated to familiarize the pilots with the varying display content that they might encounter on a descent from 15,000 ft to the outer marker. (Note that "miles" refers to nautical miles in this report.) The areas covered by the six map scales were 4 mi. (fig. 1), 8 mi., 16 mi., 32 mi. (fig 2), 64 mi., and 128 mi. (fig 3) from top to bottom and from side to side.

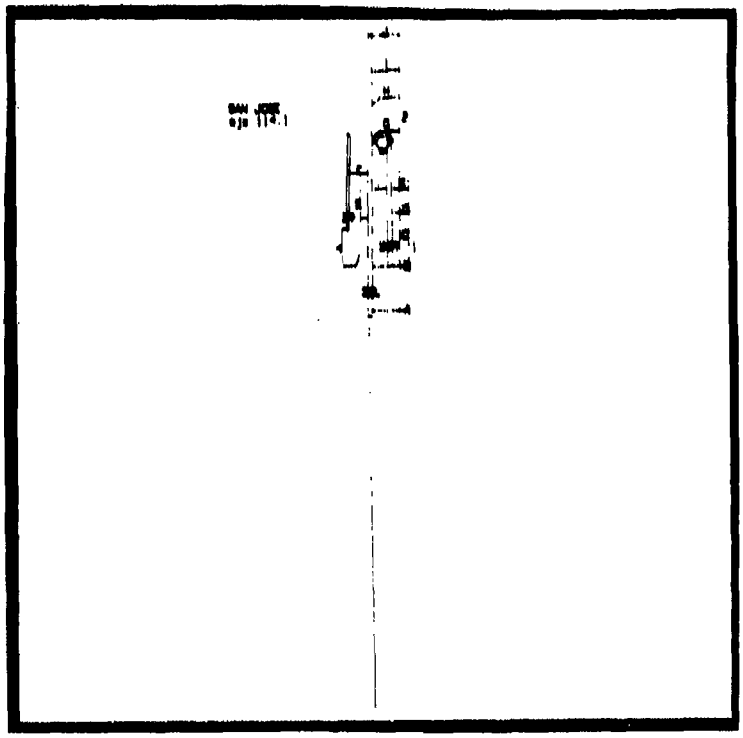


Fig. 1: Navigation display for southern approach to San Jose Municipal Airport: 4 miles full scale

Since the displays in the study were static, map scaling per se was not a variable, however, the amount and type of information that pilots felt should be displayed at different altitudes, for different map ranges, and during different phases of flight were examined. All of the maps were presented with a heading-up orientation.

Navigation display. Since a CDTI could perform multiple functions, consideration was given to providing a graphic display of routes, navigation aids, airports, and intersections to assist the pilot in placing the position of his own and other aircraft in context and to use for primary navigation.

Fig 2: Navigation display for southern approach to San Jose Municipal Airport: 32 miles full scale

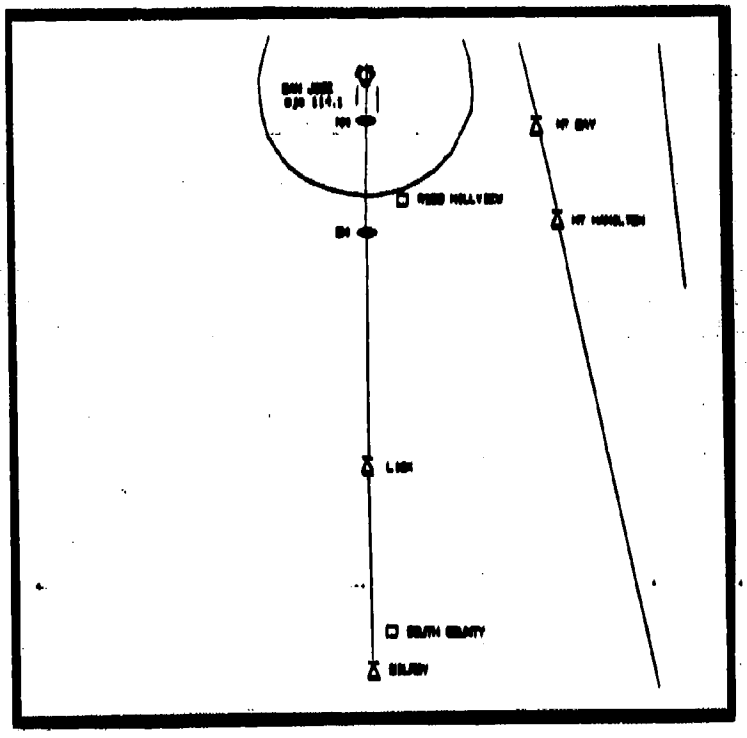
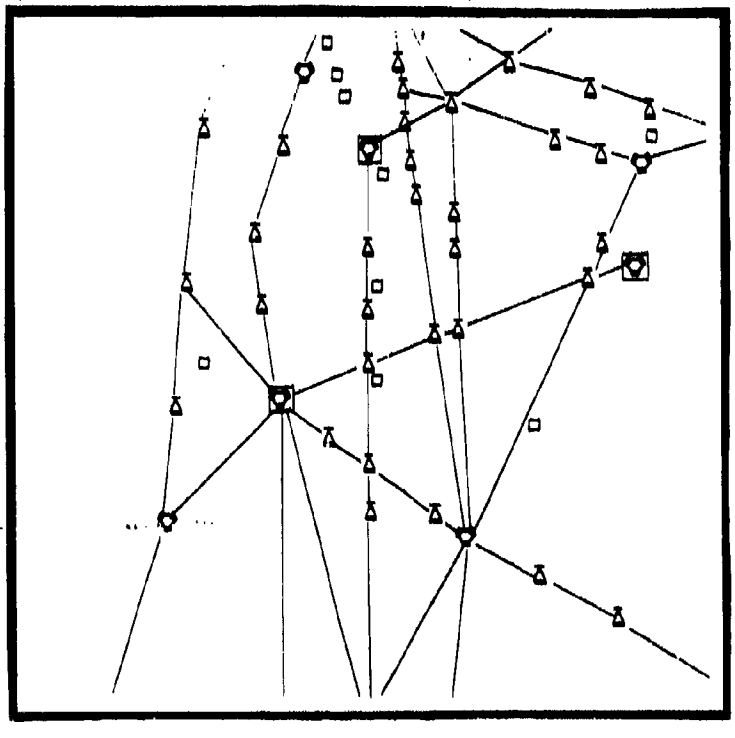


Fig. 3: Navigation display for southern approach to San Jose Municipal Airport: 128 miles full scale.



The variables investiagted included: (1) navigation routes to display - own route only, own and intersecting routes, or all routes with the range of the map; (2) symbology; (3) labeling; (4) color coding; and (5) display of current map scale and map orientation.

Terrain display. Again in order to determine whether or not several categories of information should be combined on a CDTI, several different ways to display information about terrain were presented. The variables investigated included: (1) Should terrain be displayed on a CDTI? (2) Should terrain be displayed at pilot request or automatically? (3) What obstructions should be displayed? (4) symbology; and (5) color coding. All significant man-made and natural obstructions that were within 5000 vertical feet of the simulated altitude of own aircraft were graphically superimposed on each of the maps with height labeled in feet. (fig. 4)

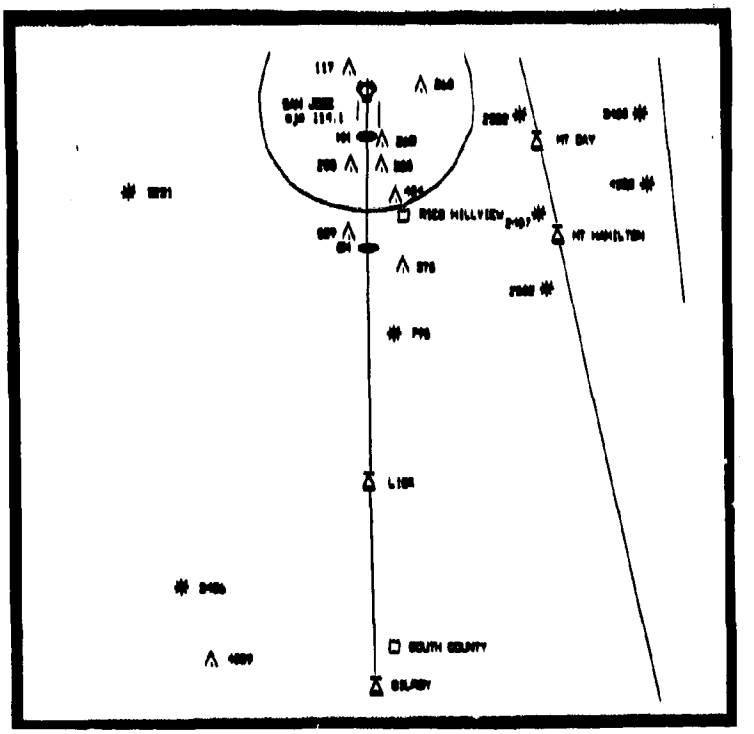


Fig. 4: Terrain symbology superimposed on a 32 mile map representing a southern approach to San Jose Airport

In addition, grid lines were superimposed on each map and the minimum safe altitude within each section was presented in hundreds of feet. This method of presentation was suggested as an alternative or addition to the display of specific terrain features. In concept, the grid display was similar to the Minimum Safe Altitude Warning (MSAW) data base currently in use by ATC with which controllers automatically receive a warning if an aircraft is in immediate jeopardy or it is predicted that it will be within 30 sec.

Weather display. A CDTI could also include a graphic display of weather as well as other types of information. The variables studied included: (1) Should weather be displayed on a CDTI? (2) Should a weather display appear at pilot request and/or automatically? (3) What information about weather should be displayed? (4) symbology; and (5) color coding. The pilots were shown three display formats depicting different amounts of information about weather: (1) the location of weather was depicted by radial lines emanating from a radar site with a different letter indicating the nature of the weather and areas of intensity. (2) Letters were coded to display the location and and nature of areas of heavy precipitation only; and (3) A random dot pattern was superimposed on the map to show location only with no indication of intensity or nature. (fig. 5)

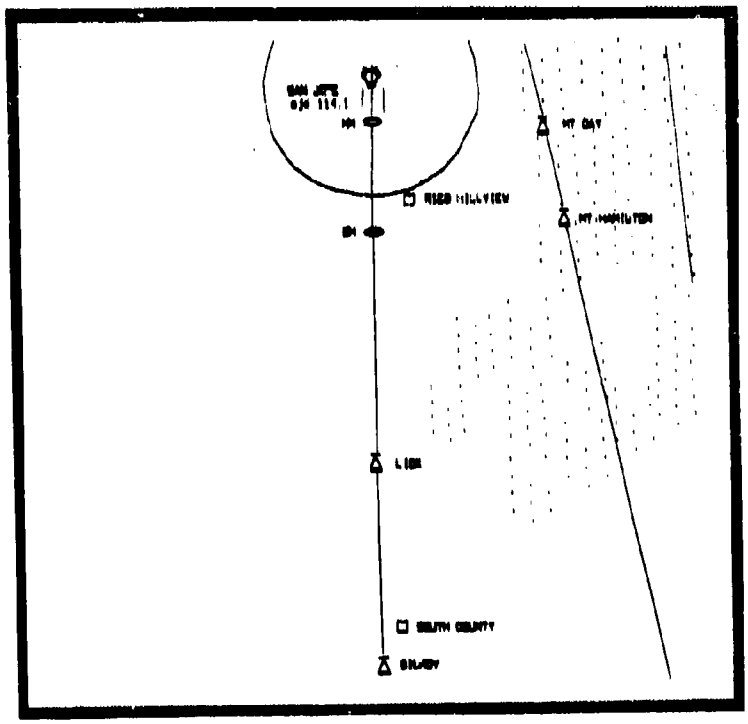


Fig. 5: Graphic display of the location of weather superimposed on a 32 mile map representing a southern approach to San Jose Municipal Airport.

Own aircraft. Different ways to represent a pilot's own aircraft were investigated with particular emphasis on: (1) symbology; (2) location of the symbol on the display; and (3) the relationship between the representation for own and other aircraft. Six symbols, including most of the symbols (fig. 6) in current use, were shown to the pilots on a single display with and without flight path predictors and in the context of a basic navigation display. Three possible vertical locations for the own aircraft symbol were shown (centered laterally): centered; offset so that 2/3 of the map was ahead; and offset so that 3/4 of the map was ahead.

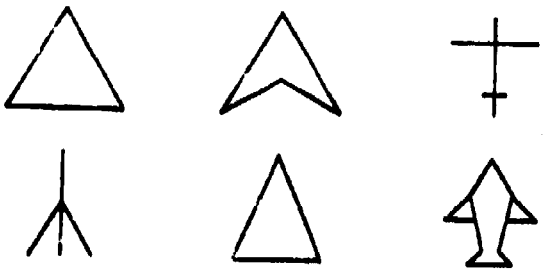


Fig. 6: Candidate symbols for own aircraft.

Traffic display. Since the primary function of a CDTI will be to provide a graphic display of adjacent traffic, a major focus of the study was to determine: (1) What proportion of other aircraft should be displayed? (2) symbology; (3) What additional information about the status of other aircraft should be displayed? and (4) How should such information be displayed (data tags, coded symbols, or data tables on a second display).

Since it is unlikely that all aircraft within the range of the map will be displayed because of clutter, the logic for displaying other aircraft was investigated. The algorithm could be related to the altitude, speed, or map scale of own aircraft, proximity to own route, or some combination.

The simplest representation of another aircraft would be a nondirectional symbol that represented its position only. This type of symbol was represented by a circle in this study. In addition, the symbol could also present information about the direction of flight by its orientation (i.e. a directional symbol). An isosceles triangle and a track line added to the circular, nondirectional symbol were suggested as two ways to display direction of flight. In addition, flight path history or "trail" was provided for half the displays. This also provided information about direction of flight by extrapolation. Each of these symbols with and without flight path history was used as the symbol for traffic in 32 and 128 mi. maps which simulated traffic densities typical of the San Jose area. All aircraft within +/-4000 ft of the pilot's own altitude were included. The same positions, altitudes and densities were used for each of the displays. An example of one such display may be seen in Figure 7.

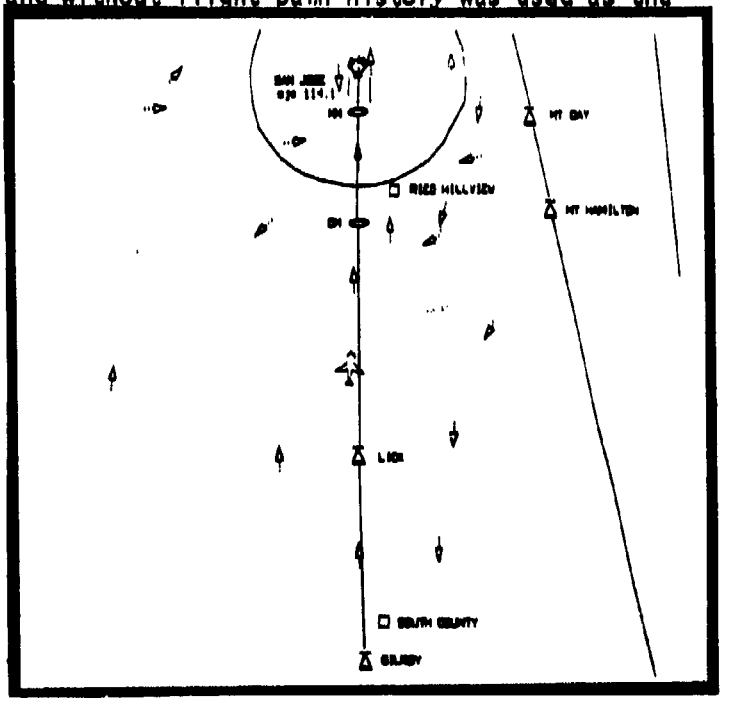


Fig. 7: Traffic display superimposed on a 32 mi. map: Directional symbols with flight path history

The shape of the symbol used to depict the position of other aircraft could be varied to display additional information about the status of other aircraft graphically. Examples of eight coding schemes were shown in which the shapes of the symbols were varied to depict ATC status (under ATC control or not) alone or ATC status and CDTI equippage and/or the relative altitude of the aircraft with respect to the altitude of a pilot's own aircraft. Two relative altitude coding concepts were shown: (1) the non-directional symbol differentially indicated whether an aircraft was at, above, or below the altitude on a pilot's own altitude; and (2) the non-directional symbol showed only whether or not an aircraft was at the same altitude as the pilot's own aircraft or not, but did not differentiate between aircraft that were above or below. An example on the most complex coding schemes may be seen in Figure 8 and in the context of a 128 mi. map in Figure 9.

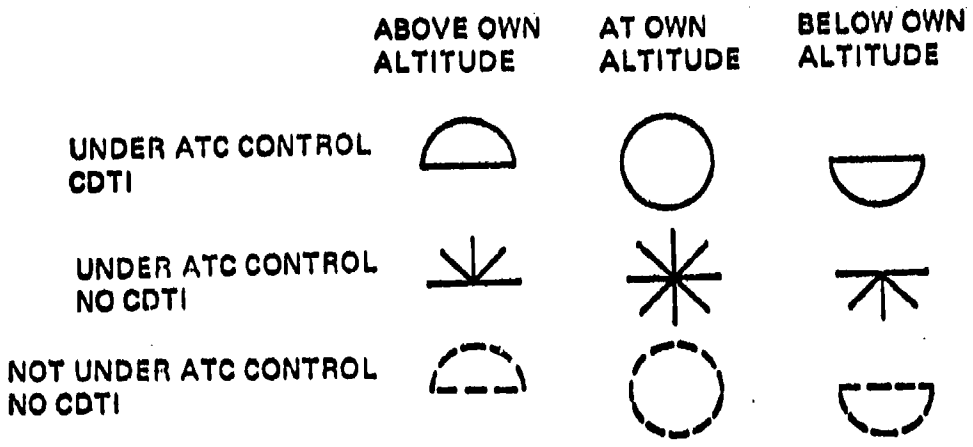


Fig 8: Non-directional symbols depicting relative altitude, ATC status and CDTI equippage

The pilots were also asked to specify what additional information they felt should be provided about the status of other aircraft by rating 18 different times individually, and to specify the preferred display format for the different types of information: (1) symbol encoding; (2) digital data blocks; (3) digital data blocks displayed at pilot request only; and (4) an alphanumeric table on a separate display.

Effectiveness of multicolor displays. Pilots were shown a cross section of the displays presented simultaneously monochromatically (green only) and with color coding (green for navigation, yellow for terrain and weather, and red for own and other aircraft) to determine whether the pilots felt that color coding was necessary.

Concluding questions. The final section involved several question about display features that the pilots had not been shown (e.g. map scaling and display size) and the pilots' evaluation of the potential usefulness of the information that a CDTI could provide and its potential impact on their workload.

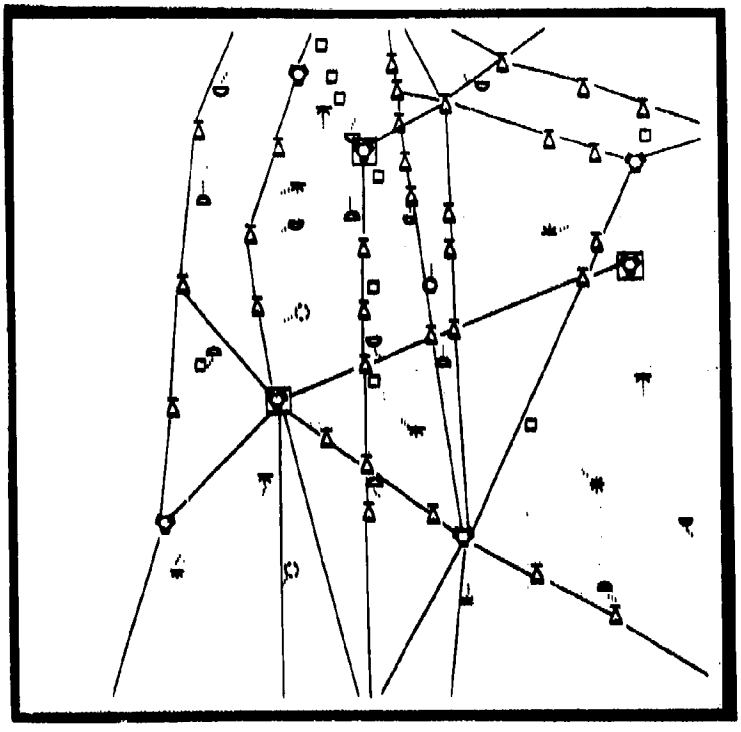


Fig. 9: Traffic display superimposed on a 128 mi. map: Non-directional symbols with flight path history coded to depict relative altitude, ATC status and CDTI equippage

RESULTS

The significance of the pilots' responses was computed by calculating the probability that one or more of the alternative choices allowed for a given question would be chosen R or more times by chance alone. The value of R was computed for p=.05 and p=.01 for the 95% and 99% confidence levels respectively (indicated by * and ** in the figures) The critical values required to assert that a significant preference was found may be seen in Table I.

In the following section, the opinions expressed by a significant number of the general aviation pilots included in this study will be summarized and their responses will be compared to those obtained from the airline pilots in the earlier study.

It is essential that any conclusions that the reader may draw from these results take into account the limitations of the research methodology: The displays were presented statically and the pilots could not interact with them or use them in an operational environment. (2) The results reflect subjective evaluations rather than objective measures of performance. (3) It was impossible to present all possible format and symbology options and combinations and thus alternatives may exist that the pilots would have selected had they been shown to them

Table 1

Minimum number of subjects selecting a specific alternative required to assert that a significant preference was evident in the responses of the 13 general aviation pilots

		Number of alternative responses possible						
		2	3	4	5	6	7	8
Level of Confidence	95%	11	9	8	7	7	6	6
	99%	12	10	9	8	7	7	7

Navigation display

The general aviation pilots gave essentially the same responses as did the airline pilots to the questions posed about the navigation portion of the display. They indicated that the proportion of routes displayed should increase as the range of the map is increased (e.g. own route only for 1-10 mi. map scales, own and intersecting routes for 10-50 mi. map scales and either own and intersecting or all routes within the range of the map for map scales greater than 50 mi.) (fig. 10) Primary navaids, intersections, and airports should be displayed symbolically. Identifying labels should be included for intermediate map scales. A digital indication of direction of flight and map scale should be provided. Airport control zones, taxiways, and TCAs should not be displayed.

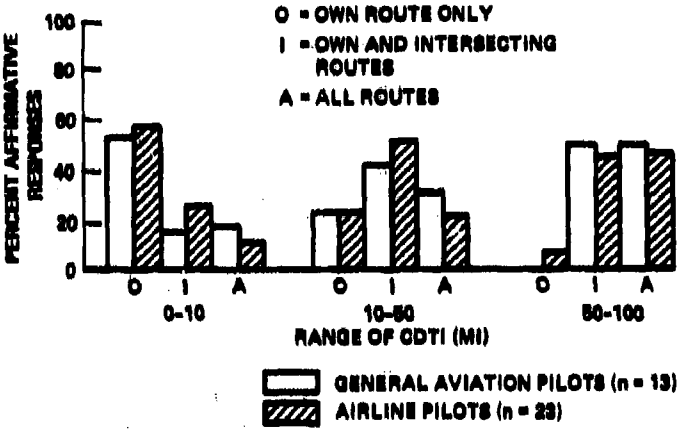


Fig. 10: Proportion of routes to display for different map scales

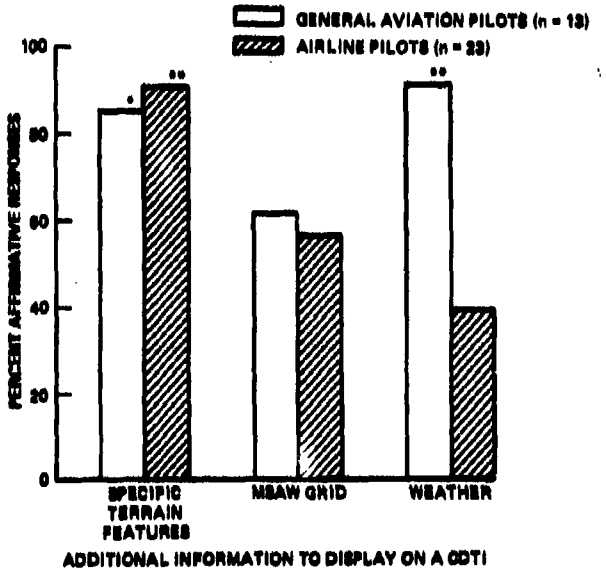


Fig. 11: Acceptability of including specific terrain features, minimun safe altitude grid and/or weather on a CDTI

Terrain display

Eighty five percent of the general aviation pilots felt that significant terrain features should be displayed by specific symbols to represent the location and heigth of individual obstructions, as compared to 91% of the airline pilots (fig. 11). More than half of both groups of pilots felt that a digital readout of the minimum safe altitude for different areas would be an acceptable alternative to a symbolic display of specific obstructions. (fig. 11) If a symbolic display is used, one or two symbols (one for man-made, ⏃, and one for natural, ✳, obstructions) would be acceptable with heigth labeled in feet.

A significant percentage of both groups of pilots felt that terrain information should be displayed automatically if an aircraft is below the minimum safe altitude and should be available at pilot request as well, regardless of its format. (fig. 12) Ninty three percent of the general aviation pilots felt that obstructions 2000 ft. or closer should be the only ones displayed (as compared to

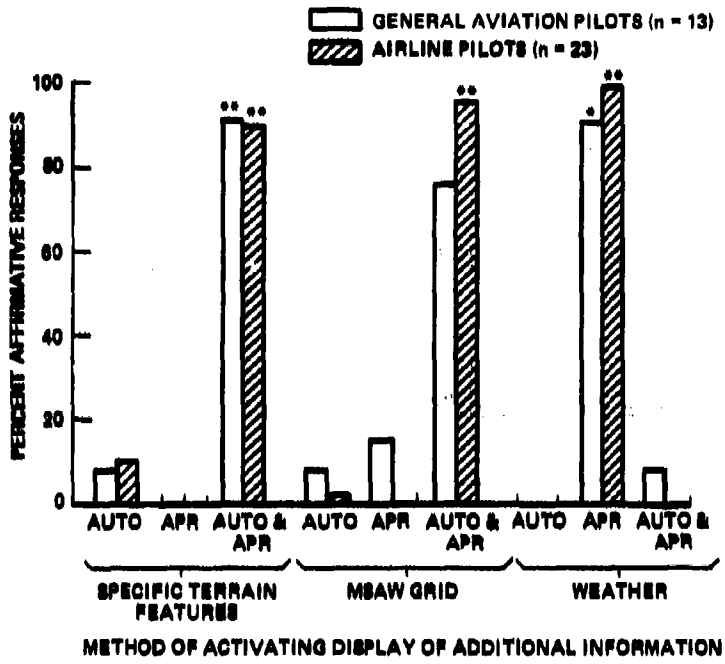


Fig. 12: Preferred method of activating a graphic display of terrain, minimum safe altitude, or weather: Automatic if conditions warrant (Auto), at pilot request (APR), or both (Auto & APR)

59% of the airline pilots, many of whom felt that the algorithm for displaying terrain should vary with the altitude of the pilot's own aircraft.

Weather display

The general aviation pilots differed from the airline pilots with respect to their willingness to accept a graphic display of weather on a CDTI. Ninty two percent of the general aviation pilots felt that weather should be included, whereas the majority of the airline pilots preferred to have a separate display dedicated to weather. (fig. 11) If weather is to be included on a CDTI, a significant number of both groups felt that the display should include information about the nature and intensity of the weather as well as its location, although neither group preferred any one of the display formats shown to them over the others. Ninty two percent of the general aviation pilots, and all of the airline pilots, felt that any display of weather on a CDTI should be initiated by the pilot and should not appear automatically. (fig. 12)

Symbology for own aircraft

All but one of the general aviation pilots selected the most airplane-like symbol to represent their own aircraft. (fig. 13b) The airline pilots had selected this symbol as their second choice and the chevron-shaped symbol as their first choice. Triangles (isoceles or equilateral) were considered to be unacceptable by more than 70% of the general aviation pilots (fig. 13a) and was selected by none of them and by few of the airline pilots as their first choice. Although more of the general aviation pilots found the stick figure symbols to be acceptable than did the airline pilots, none of them selected either of the stick-figure symbols as their first choice and few of

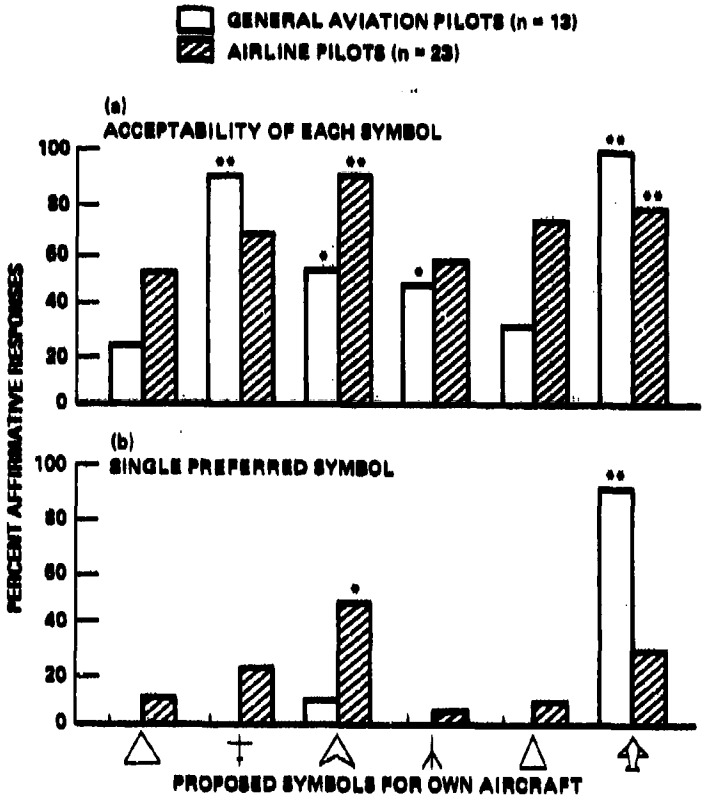


Fig. 13: Pilot response to symbols proposed for their own aircraft

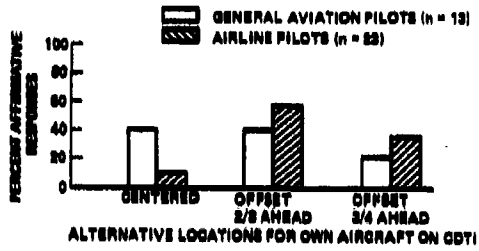


Fig. 14: Preferred location for own aircraft symbol on CDTI

the airline pilots did. Both groups of pilots felt that the symbol for their own aircraft should be clearly differentiated from the symbol(s) for other aircraft by size, shape, and/or color.

There was no concensus among the general aviation pilots about the optimal location for their own aircraft symbol on the display. Although 92% of the airline pilots felt that the symbol should be offset so that proportionally more of the map was ahead, (fig. 14) 38% of the general aviation pilots felt that it should be centered on the CDTI.

Traffic display

Although none of the airline pilots wanted all aircraft within the range of the CDTI to be displayed routinely, 30% of the general aviation pilots selected this option. If, however, the proportion of other aircraft displayed was to be limited, 30% of the general aviation and 48% of the airline pilots felt that only those aircraft on own and intersecting routes should be displayed, and the majority of both groups felt that own aircraft speed, altitude and map scale should be included in any algorithm for limiting the display of other aircraft. Both groups of pilots agreed that limiting the proportion of other aircraft displayed to those within +/- 2000 ft. of their own altitude and adjacent laterally would be an acceptable range within which to display all aircraft for performing the tasks of fine tuning and merging. A wider range (e. g. all aircraft within 2-4000 ft. above and below own altitude) might be displayed for monitoring the traffic situation.

In general, the general aviation pilots found the non-directional symbols with or without flight path history to be less acceptable than did the airline pilots. (fig. 15) More than 90% of both groups found the directional

triangular symbol with flight path history to be acceptable.

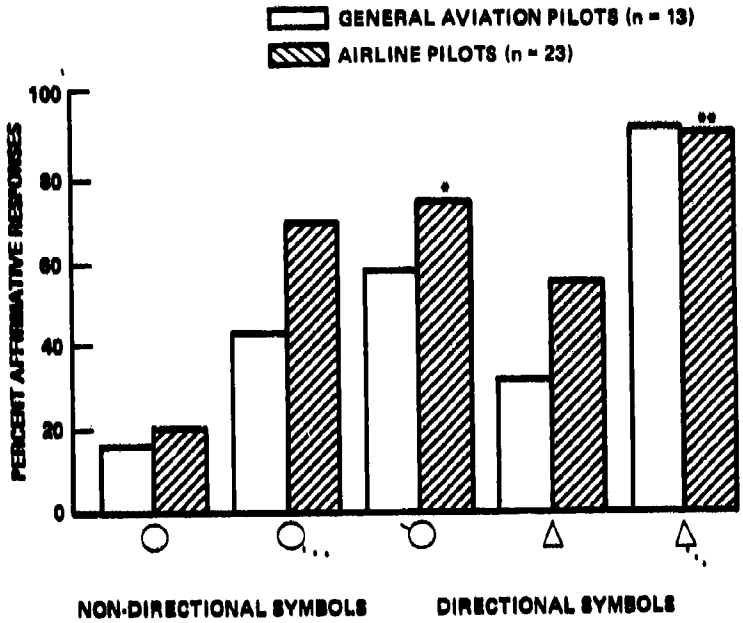


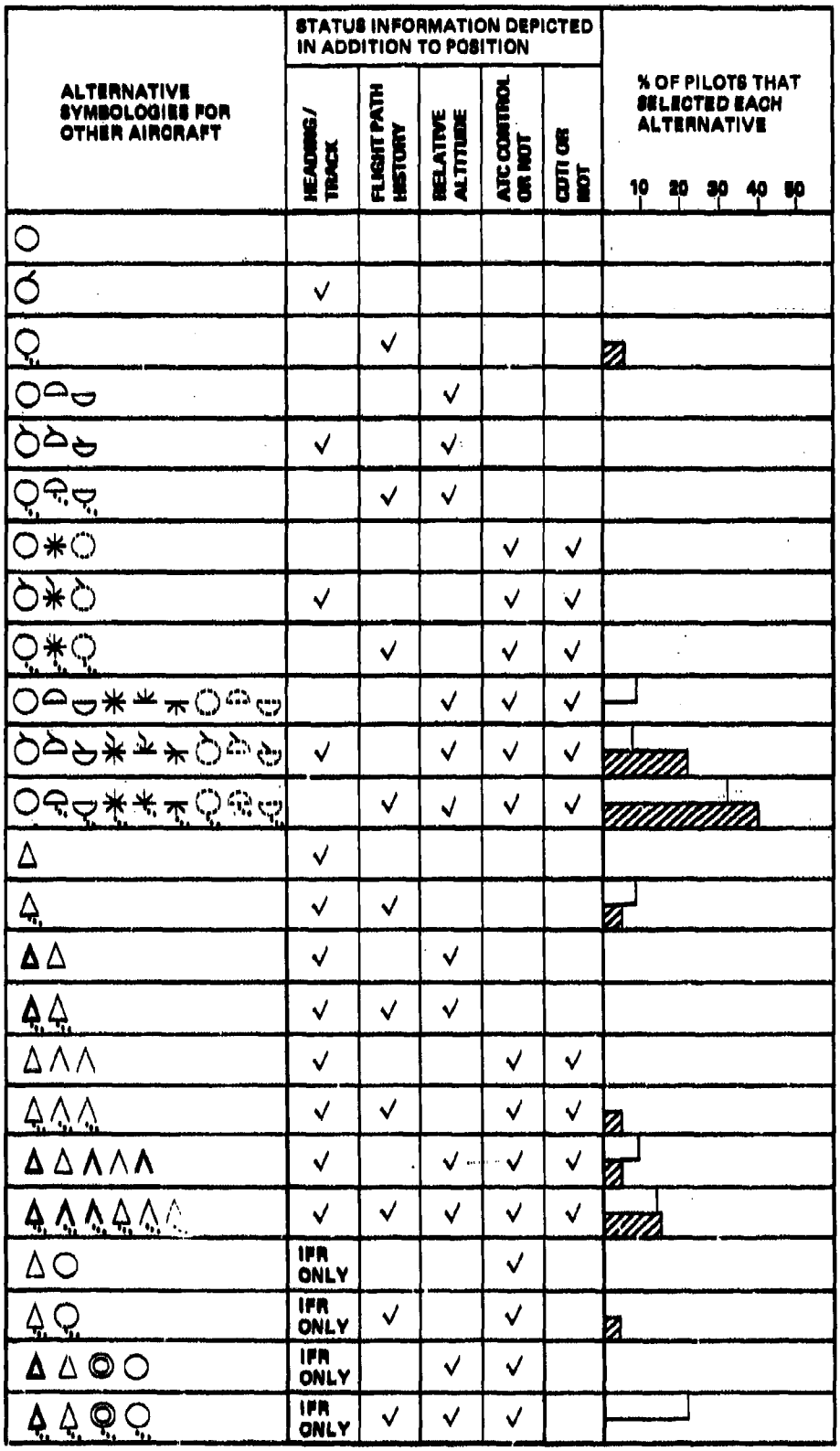
Fig. 15: Pilot rated acceptability of the basic symbols suggested for other aircraft

Both groups of pilots favored some form of symbol shape encoding to depict the relative altitude and ATC status of another aircraft. Of the coding schemes suggested, both groups of pilots found the relative altitude encoding scheme that indicated whether another aircraft was at, above, or below own altitude to be more acceptable than a coding scheme that simply differentiated aircraft at their own altitude from those that were not. Although both groups had preferred the triangular symbol to the circular non-directional symbol because it depicted direction of flight, they felt that information about relative altitude (at, above, or below) was so important that they were willing to forego information about direction of flight in order to obtain it. When asked to select the one symbol or set of symbols that they preferred from the 24 alternatives presented (fig. 16), more than 70% of both groups of pilots selected a symbol that included flight path history, more than 50% selected symbols that depicted direction of flight, 82% of the airline pilots and 93% of the general aviation pilots selected symbols that depicted relative altitude and 92% of both groups selected a symbol set that included information about ATC status, although only one pilot from the two groups selected a coding scheme that depicted only ATC status. None of the pilots selected a symbol for other aircraft that depicted position only. All of them wanted the symbol to convey at least some additonal information.

When asked to evaluate the additional types of information about the status of other aircraft that should be available on a CDTI, a significant number of pilots in both groups selected altitude, flight path history, and emergency status as being particularly important from the list of options. (fig. 17) In addition, more than 70% of the general aviation pilots felt that direction of vertical flight, ground track and speed, relative altitude, and landing sequence number should be displayed. The airline pilots had felt that information about weight class was preferable to information about aircraft type, and a much greater percentage of the airline pilots (83% vs. 46%) felt that knowing whether another aircraft was IFR or VFR was necessary.

Both groups of pilots felt that some of the above information could be coded into the symbol for other aircraft. Although both groups of pilots felt

Fig. 16: Pilot preferred symbology for other aircraft



that such information as altitude, ground speed, weight class or aircraft type, I.D., vertical speed, and destination should be displayed digitally in a data block beside the symbol for another aircraft, few pilots in either group wanted such information displayed at all times for all aircraft. Rather, they wanted to request such information at their discretion for specific aircraft. The majority of both groups preferred the option of a touch sensitive display as a means of requesting the display of a digital data block for another aircraft. Few pilots selected the option of an additional display for digital information about other aircraft.

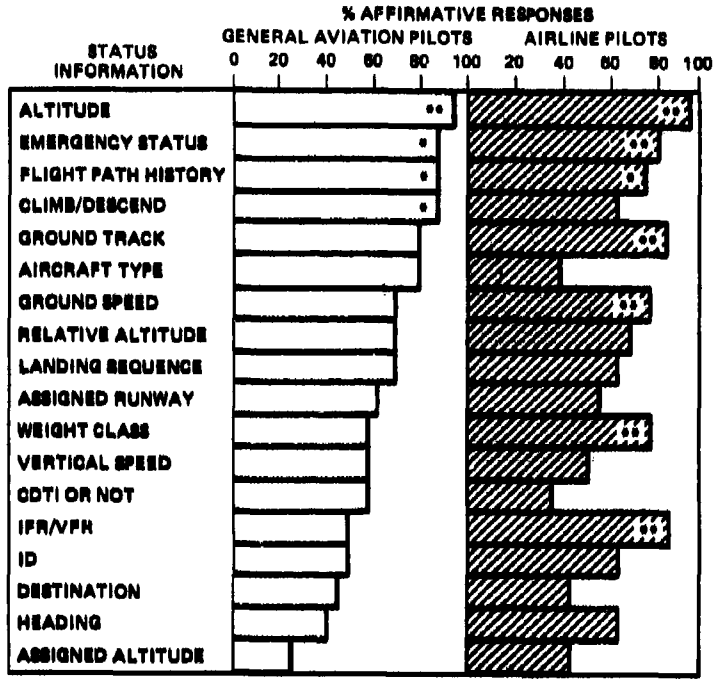


Fig. 17: Pilot rated necessity for different types of information on the status of other aircraft.

Use of color

All of the pilots in both groups felt that color coding was essential to differentiate among categories of information and to improve their speed and accuracy of recognition. More than 85% of both groups felt that color encoding would assist them in determining the position of other aircraft and 91% of the airline and 69% of the general aviation pilots felt that color encoding would assist them in maintaining separation from other aircraft alone or in combination with symbol encoding.

Display size

The general aviation pilots were willing to accept a much smaller display size than were the airline pilots. (fig. 18) Whereas only one airline pilot was willing to accept a CDTI less than 7 in. in diameter, six out of 13 general aviation pilots felt 5 in. would be the minimum acceptable size.

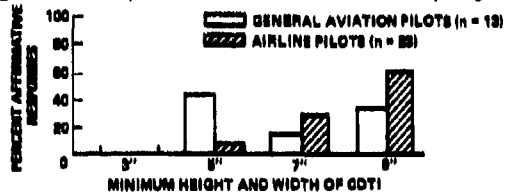
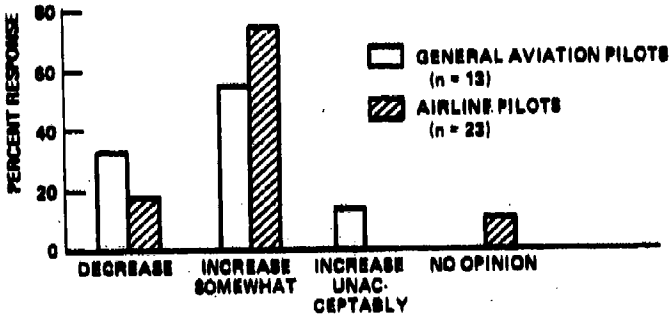


Fig. 18: Minimum acceptable size for a CDTI

Opinions about potential impact of CDTI

Only one general aviation pilot and none of the airline pilots felt that the additional task of monitoring a CDTI would increase their workload to an unacceptable level. (fig. 19a) Fifty eight percent of the general aviation pilots and 74% of the airline pilots felt that their workload would increase somewhat with the addition of a CDTI. A number of the general aviation pilots expressed their concern about the additional

(a) Impact of CDTI on workload



(b) Will CDTI provide useful information about the positions and intentions of other aircraft?

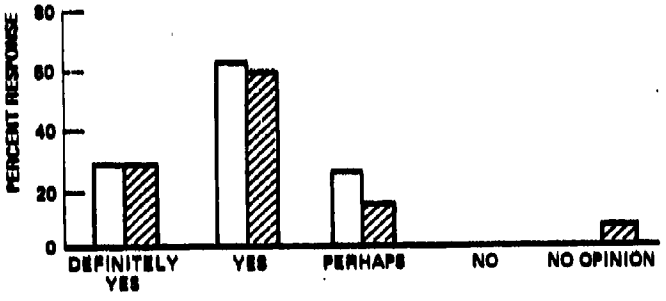


Fig. 19: Pilots' responses about the potential impact on CDTI

responsibility of monitoring a CDTI when flying without a co-pilot under IMC. Despite this concern, 31% of the general aviation pilots and 18% of the airline pilots felt that their workload might actually decrease with the addition of a CDTI.

All but three of the general aviation pilots and four of the airline pilots felt that a CDTI would provide useful information about the position and intention of another aircraft. (fig. 19b) Three additional pilots in each group felt that such a display would be useful if, and only if, it was a fundamental display that combined multiple categories of information and functions. In addition, several general aviation pilots exprassed their concern that implementing CDTI in the ATC system could mean the "end of VFR flight".

Even though both groups of pilots expressed concern about display clutter, they both felt that it would be necessary to have a great deal of information available for display on a CDTI. Their solution to this inconsistency was to suggest that the pilot should be allowed to control the amount, type, and complexity of information displayed at any time.

It is possible that the pilots may revise their opinions about the amount and complexity of information to be displayed, and about symbology and format, once they have used these displays in a dynamic simulation. For this reason, it must be emphasized again that the results of this study, and the earlier study, reflect pilot opinion only and that different results might be obtained once the pilots have used the different displays in simulated flight.

One of the most interesting results of the present study was the degree to which the two groups of pilots agreed about the optimal content, format and symbology on a CDTI. This occurred even though the groups differed with respect to age, flight experience, and operational environment. Differences that were found appeared to be related to the general aviation pilots' concern with cost and the additional load that such a modification would impose on a single pilot

particularly under adverse weather conditions. It appears that there is considerable uniformity among the group of pilots sampled about the information that they felt should be displayed on a CDTI (with the exception of including a display of weather), the format in which it should be presented, and the symbology to be used (with the exception of the symbol selected for own aircraft). These results were obtained from 36 pilots who differed from each other with respect to type of aircraft flown, cockpit environment (1, 2, or 3 man), air traffic control environment typically flown in, years of experience, and reason for flying. It will be interesting to see whether this commonality of opinion persists into the dynamic simulation environment which will, of course, be a much more valid test of the concepts presently under consideration.

REFERENCES

1. Cockpit Displayed Traffic Information Study (FAA-EM-78-18) Washington, D.C.: U. S. Government Printing Office, September 1977. (Prepared by the Boeing Commercial Airplane Company.)

2. Hart, S. G., McPherson, D., Kreifeldt, J. and Wempe, T. E. Multiple Curved Descending Approaches and the Air Traffic Control Problem. (NASA TM78, 430) Washington, D. C.: U.S. Government Printing Office, 1977.

3. Connelly, M. E. Simulation Studies of Airborne Traffic Situation Display Applications - Final Report. (MIT Report ESL-R-751) Cambridge, Mass.: Massachusetts Institute of Technology, May 1977.

4. Hart, S. G. & Wempe, T. E. Cockpit Display of Traffic Information: Air Carrier Pilots' Opinions about Content, Symbology and Format. (NASA TM in preparation)

5. Kaiser, D. H. Design and use of a human engineered system for the easy generation of static images. Paper presented at the Meeting of the Society for Computer Simulation, Graphics, and Modeling. Monterey, Ca, April 10, 1979.

PERCEPTION OF AIRCRAFT SEPARATION WITH VARIOUS SYMBOLS ON A COCKPIT DISPLAY OF TRAFFIC INFORMATION

by
Everett Palmer, Daniel Baty, and Sharon O'Connor
Man-Vehicle Systems Research Division
Ames Research Center, NASA, Moffett Field, California 94035

ABSTRACT

Perception of motion and aircraft separation on a cockpit display of traffic information (CDTI) may be affected by many different display elements such as: information content of aircraft predictors and history, number and type of display background elements, map orientation, map scale and update rate. An experimental approach is described for quickly evaluating the relative importance of many of these display variables. The procedure requires that the pilot judge whether an intruder aircraft will pass in front of or in back of his own aircraft based on a short look at the encounter situation at some time prior to the time of closest approach.

The results of four experiments, each of which dealt with a different set of display variables, are described. Display variables were horizontal flight path predictor types and history for both own-ship and intruder. Pilots made fewer judgment errors when they had predictive information, especially with the predictor curved proportionally to turn rate.

Plans for the experimental series, which will continue, are briefly explained. A small subset of display elements will be recommended for further evaluation in a full mission simulation.

INTRODUCTION

The concept of displaying pertinent traffic information in the cockpit has been proposed for many years. Only recently, however, have crowded airspace conditions added pressure to finding new solutions for safely maintaining less separation distance between aircraft (ref. 1). Three proposed methods of displaying such traffic information are: (1) a display dedicated only to traffic information; (2) adding traffic information to the weather radar display; and (3) integrating traffic information with an electronic map display. In the third option, the CDTI would also serve as the pilot's primary display of navigation and guidance information. One's own position and direction of travel with respect to area navigation routes and terrain features would be indicated by a "heading up" or a "track up" moving map display. The display computer would continually translate and rotate this map so that current aircraft location would be represented by a fixed aircraft symbol and current value of aircraft heading or track angle would be displayed at the top of the display. Objects on the display would move down the display at a rate proportional to aircraft movement over the ground. When own-ship

turns, all objects on the map would rotate about the fixed aircraft symbol. Symbols showing the location of other air traffic would move with respect to both ground referenced objects and own-ship.

Perceptually, a CDTI presents a more complicated display than an air traffic controller's (ATC) display. On an ATC display, the map is fixed and all of the aircraft symbols move with respect to that fixed background; in the CDTI, the pilot is presented with a continuously moving background as well as moving aircraft. A parallel can be drawn between perceptual problems with conventional shipboard radar and potential CDTI problems. Perceptual errors caused by misleading, apparent motion of other ships, due to own-ships rotation, has occasionally led to what are referred to as "radar assisted collisions" (ref. 2). Sophisticated collision-avoidance systems are now being developed for ships to improve the raw radar display in order to eliminate those perceptual problems. One goal of the research described in this paper is to specify relevant information for a CDTI so that initial users will not suffer from problems analogous to those of early conventional radar displays in ships.

The first step taken toward developing a clear and easy to use generic CDTI was to develop an exhaustive list of questions regarding the CDTI concept. The questions pertinent to the generic CDTI issue were examined and an experimental approach to answering each question was determined. The questions to be answered dealt with the display symbols and density, factors altering perception of motion, pilot communication with the display, CDTI interaction with other systems, and display location. This study deals with some basic factors affecting pilot perception of motion and traffic separation. Pilots made judgments while monitoring a dynamic CDTI display. Errors in judgment were recorded to determine how accurately pilots could predict the future separation between their own aircraft and an intruder aircraft. Different combinations of displayed information and task difficulty were chosen for each of four experiments. The method developed here will also be used to examine additional display variables. In other proposed experiments, pilots will manually control the aircraft simulator for such tasks as in-trail spacing, merging, and collision avoidance maneuvering using the display symbols developed in these experiments.

METHOD

Display Hardware

The CDTIs were two 7- by 7-in. CRTs located directly below the attitude indicators of the pilot and copilot in a fixed-base cockpit simulator. The center of the display was 25° below the horizontal and was 35 in. from the pilot's eye reference point. The display symbols were generated by a general purpose, stroke-writing computer graphics system. The green phosphor on the CRT's left no noticeable after glow.

Display Symbology

Figure 1 shows four of the display formats used in the experiments. The following display symbols were constant in this series of experiments. The

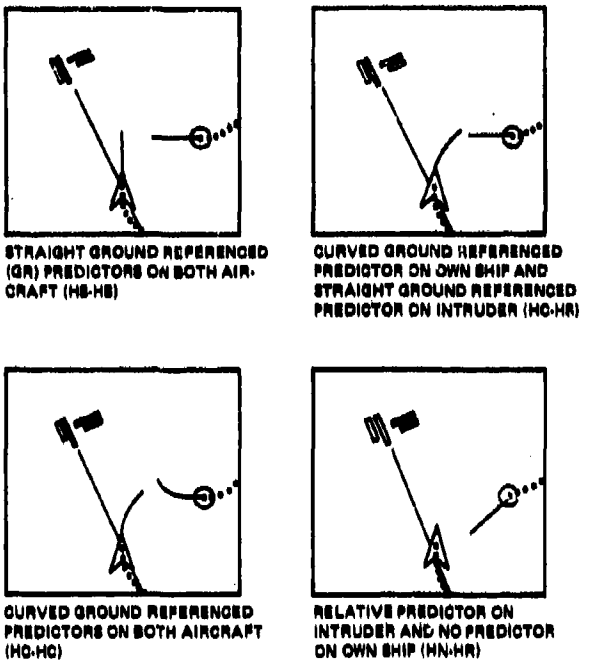


Figure 1. Four of the display formats used in the experiment. The following parameters were not changed: relative bearing at encounter, ϕ_r = 0°; own-ship speed = 200 knots; background present; update every 4 s.; 32 s. of ground-referenced history; map track up; and width of map = 10 n. mi. of terrain.

top point of the chevron symbol indicated the present position of own-ship. Present position of the intruder was indicated by a dot in the center of a circular symbol. These symbols were preferred by the most pilots in Hart's study (ref. 3) of pilot opinions on various types of possible CDTI symbols. An RNAV route and runway symbol provided ground objects for background. The width of the terrain displayed on the map display was always 10 n. mi. With this map scale, which seems reasonable for terminal area operations, 1 n. mi. on the ground equals 0.5 in. on the display. The display was rotated so that own-ship's track was at the top of the map. Own-ship position and all intruder in formation were updated every 4 sec. No sensor noise or filter lag was simulated for these tests.

The independent display variables included the presence or absence of ground referenced history and the type of predictor on own-ship and intruder. Table 1 shows the various combinations of display symbols used in the four experiments and defines the kind of history and predictor information displayed. These displays are discussed in more detail in the description of the individual experiments.

TABLE 1

COMBINATIONS OF DISPLAY SYMBOLS USED IN EXPERIMENTS 1 THROUGH 4

Own-ship		Intruder*		Code	Experiment			
History	Predictor	History	Predictor		1	2	3	4
None	None	None	None	NN-NN	X			
History	None	History	None	HN-HN	X	X		
None	Straight	None	Straight	NS-NS	X	X		
History	Straight	History	Straight	HS-HS	X	X	X	X
History	None	History	Relative	HN-HR			X	
History	Curved	History	Straight	HC-HS				X
History	Curved	History	Curved	HC-HC			X	X

*Definitions

None: symbol was not present.

History: Every 4 sec. a new history dot was added marking the current ground position of the aircraft. A total of 8 dots or 32 sec. of history were displayed.

Straight: The end of the vector predicts the position of the aircraft over the ground in 32 sec., assuming the aircraft maintains its current ground track

Curved: The end of the vector predicts the position of the aircraft over the ground in 32 sec., assuming the aircraft maintains its current turn rate.

Relative: The end of the vector predicts the intruders position relative to own-ship in 32 sec., assuming both aircraft maintain their current ground track.

Encounter Variables

Figure 2 shows the eight parameters that were used to specify an encounter between own-ship and an intruder. The separation distance at the point of encounter, R, was an independent variable. In these experiments, R was either 3,000, 6,000, or 9,000 ft. There were no encounters that would result in a collision. For each display condition, the pilots monitored 24 different encounters before progressing to the next display condition. In 12 of those encounters the intruder would ultimately pass in front of own-ship. Figure 3 depicts these 12 encounters and their parameters as they would appear using the curved ground referenced predictor and history. The remaining 12 encounters differed from these only in that the sign of the separation parameter was reversed. Note that in 12 encounters both aircraft were going straight and in the remaining 12 one or both aircraft was turning. During the experiment, the order of presentation of the 24 encounters was randomized. In addition, whether the pilot saw the encounter depicted in figure 3 or its mirror image was also randomized.

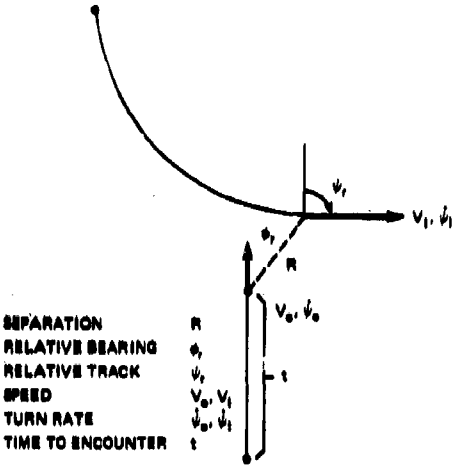


Figure 2. Parameters used to specify an encounter between own-ship and an intruder.

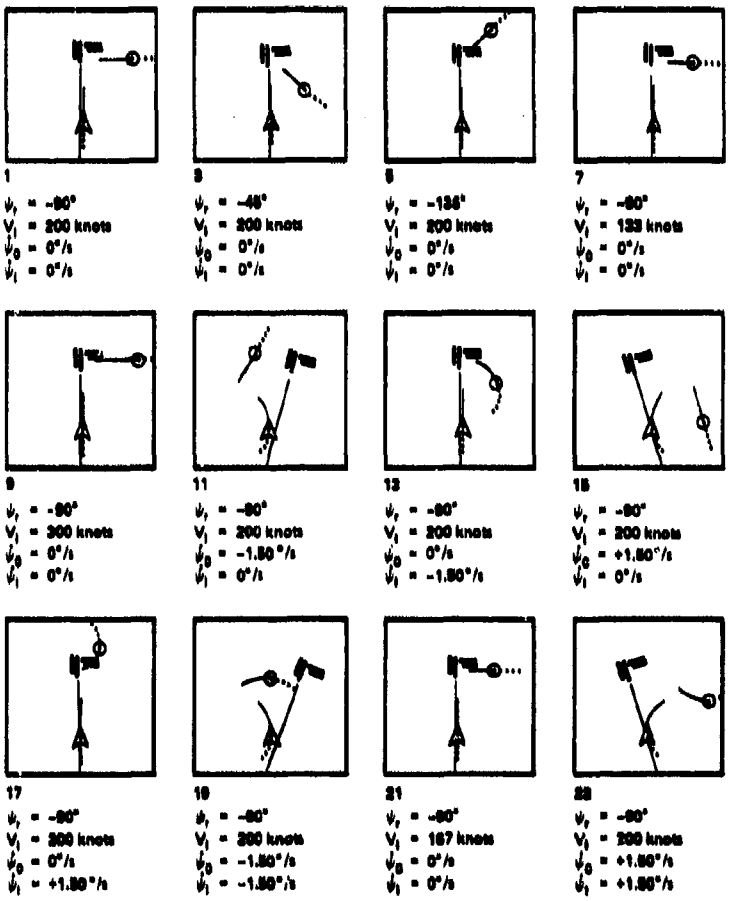


Figure 3. Twelve of the 24 encounters in which the intruder passed in front if own-ship. All displays show ground-referenced predictors. The miss distance is 3,000 ft. Relative bearing at encounter is $\phi_r = 0°$ and own-ship speed is V_O = 200 knots.

Task

The pilot's task was to monitor the CDTI and predict whether an intruder aircraft would pass in front of or in back of his aircraft. Each trial was initiated by the pilot pushing a start button. After two display updates (8 sec), the intruder appeared on the CDTI with a position, velocity, track angle, and turn rate calculated so that the intruder would be either directly in front of or in back of own-ship in 60 sec. This point is referred to as the encounter point in this paper. It is not necessarily the closest point of approach. After four additional updates (16 sec) the CDTI was blanked and replaced by a message asking whether the intruder was going to pass in front or in back of own-ship. The pilot indicated his decision by pushing a hand-held switch either forward for "in front" or back for "in back." The words "IN FRONT" or "IN BACK" then appeared indicating the correct judgment.

Subjects

Twenty-four airline pilots served as paid subjects in these experiments. They were selected from a pool of pilots who have volunteered to participate as test subjects in Ames' experiments. They were not informed of the content of the experiments until they arrived at Ames. Prior to the beginning of the study, the pilots were asked a few general questions about their attitude toward the CDTI concept and about their flying background. Sixteen of the pilots were familiar with the concept of CDTI and all of the pilots felt that the addition of traffic information in the cockpit would be desirable. Nine of the pilots were captains and 15 were first officers. The pilots' accumulated flying time ranged from 2,500 to 30,000 hrs. All but six of the pilots had military flight experience. Seven pilots were currently flying B-727's or B-737's; 5 were flying B-707's or DC-8's; and 12 were flying B-747's, DC-10's or L-1011's. All but one of the pilots had previously participated in other experiments at Ames.

Approximately 1 hr was spent describing the task and training the pilots on how to interpret the different CDTI displays before each experiment. At the conclusion of the experiment each pilot was given a four-page questionaire in which they were asked to evaluate the CDTI displays.

EXPERIMENT 1

The primary objective of experiment 1 was to determine what form of pilot response to use in later experiments. A secondary objective was to compare displays with and without straight ground referenced predictors and with and without history.

Independent Variables — Response

There were two response types. In the first, referred to here as "window" tasks, the pilots viewed an encounter for 16 sec. After 16 sec they were asked to make a decision and to respond accordingly. With this response type, two time windows were used. In the first, the intruder appeared 60 sec before encounter and in the second the intruder appeared 45 sec before the encounter. Pilot response always occurred after the display was blanked.

In the second response type, the intruder appeared 60 sec. before encounter and disappeared after 45 sec. Instead of responding when the intruder disappeared, the pilots were instructed to respond as quickly and as accurately as possible at any time after the appearance of the intruder. If they did not respond after 45 sec, the display was blanked and a message appeared asking them to respond. It was hoped that the additional data on decision time provided by this task would allow both finer discriminations to be made between displays as well as provide a measure of pilot confidence. This task is referred to as the "decision time" task.

Independent Variables — Displays

The four display variables were ground-referenced history (ON and OFF) and straight ground-referenced predictor (ON and OFF). Ground-referenced history showed the aircraft's path over the ground. It was displayed by each aircraft "dropping" a dot every 4 sec. Eight dots (32 sec of history) were displayed. These history dots translated and rotated with the terrain symbols. The straight ground-referenced predictor was a straight-line extrapolation of each aircraft's current velocity vector. It depicted where an aircraft would be relative to the ground in 32 sec, assuming that the aircraft did not turn. Note that whenever an aircraft is turning, this prediction will be in error. The magnitude of the miss distance was always 3,000 ft in the first experiment.

Experimental Design

Each of the four pilots made judgments using all four displays and all three response tasks. One days 1 and 3, pilots performed the "decision time" task. Mean error rates and mean decision times were recorded. On days 2 and 4 the pilots performed the two "window" tasks with each display conditions. Mean error rates were recorded. Pilots were run in pairs. For each task and pilot pair, the display order was randomized.

Results

Tables 2 and 3 show the percentage of errors made with each display for both straight and curved encounters for the two window tasks. The data are averaged over pilots and replications. The number of trials per cell ranged from 85 to 96. Some of the values were less than the nominal value of 96 (4-pilots × 12-encounters × 2-replications) because data from some trials were lost due to a computer hardware failure.

Comparing the error data in tables 2 and 3, it is clear that it was more difficult for the pilots to make correct judgments when the viewing time window was farther from the encounter (table 2) than it was closer (table 3). Both sets of data show similar trends. As expected, more errors were made on curved encounters, when one or both aircraft was turning, than on straight encounters when neither aircraft was turning. On the straight encounters, the straight predictors improved performance, but the addition of history did not lower the error rate. On the curved encounters, neither straight predictors nor history helped. Recall that the end of the straight predictor only correctly predicted where the aircraft would be in 32 sec when the aircraft was going straight.

TABLE 2

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR THE 60-SEC WINDOW TASK FOR FOUR DISPLAYS AND STRAIGHT AND CURVED ENCOUNTERS. THE DATA IN EACH CELL ARE THE AVERAGES FOR FOUR PILOTS, 12 ENCOUNTERS, AND TWO REPLICATIONS. THE MAGNITUDE OF THE MISS DISTANCE WAS ALWAYS 3,000 FT

Display Conditions			Encounter Types	
History	Predictor	Code	Straight	Curved
None	None	NN-NN	33 ±2	45 ±6
History	None	HN-HN	36 ±5	39 ±2
None	Straight	NS-NS	14 ±2	39 ±4
History	Straight	HS-HS	22 ±7	35 ±1

TABLE 3

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR THE 45-SEC WINDOW TASK FOR FOUR DISPLAYS AND STRAIGHT AND CURVED ENCOUNTERS. THE DATA IN EACH CELL ARE THE AVERAGES FOR FOUR PILOTS, 12 ENCOUNTERS, AND TWO REPLICATIONS. THE MAGNITUDE OF THE MISS DISTANCE WAS ALWAYS 3,000 FT

Display Conditions			Encounter Types	
History	Predictor	Code	Straight	Curved
None	None	NN-HH	24 ±2	31 ±8
History	None	HN-HN	27 ±2	30 ±3
None	Straight	HS-HS	6 ±6	31 ±6
History	Straight	HS-HS	6 ±3	34 ±6

Table 4 shows the percentage of error and mean decision times for the decision time task. The percentage error data are similar to the data from the 45-sec window task. The decision time data also indicate that even though the pilots observed the display for 2 or 3 times longer than the 16 sec allowed in the window task, their error rates were very similar to those for the 45-sec window task. The display with straight vectors (HS-HS) on the straight encounters resulted in decision times that were 8 sec shorter on the average.

TABLE 4

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR THE DECISION TIME TASK. THE NUMBERS IN THE PARENTHESIS INDICATE THE MEAN RESPONSE TIME. THE DATA IN EACH CELL ARE THE AVERAGES FOR FOUR PILOTS, 12 ENCOUNTERS, AND TWO REPLICATIONS

Display Conditions			Encounter Types	
History	Predictor	Code	Straight	Curved
None	None	NN-NN	17 ±2 (40 ±1s)	26 ±5 (42 ±1s)
History	None	HN-HN	18 ±5 (39 ±1s)	25 ±5 (42 ±2s)
None	Straight	NS-NS	2 ±1 (32 ±2s)	21 ±4 (37 ±2s)
History	Straight	HS-HS	7 ±3 (31 ±1s)	24 ±5 (38 ±1s)

An analysis of variance (ANOVA) of the error rate on the window tasks indicated a significant difference between type of display ($F(3,9) = 4.06$, $p < 0.05$). No significant difference was found for days, windows, or any interactions. Analysis of the decision time data indicated a significant difference among displays ($F(3,9) = 15.206$, $p < 0.001$); other main effects and interactions were found to be nonsignificant.

The responses to the debriefing questionnaire indicated that all four pilots preferred to have both history and straight predictor (HS-HS). All four also thought this display condition was easiest and fastest to interpret. Their comments indicated that they felt that the history and straight vector symbols made the task easier. One pilot commented that this display "did not require as much concentration." The condition that was liked the least and considered the most difficult to interpret was the display with neither history nor predictor (NN-NN). A summary of the remaining questionnaire items that are of a more general nature will be discussed after the description of experiment 4.

Based on the following considerations, it was decided to use the window task in the remaining three experiments: (1) percentage error performance was only slightly better on the decision time task than on the 45-sec window task even though the pilots had more time to look at the display; (2) the window task has only a single dependent measure which avoids the problems of tasks that have a speed-accuracy trade-off; and (3) the window task took less time to run. The 60-sec time window was chosen in order to provide a task with an error rate close to 25%.

EXPERIMENT 2

The objective of experiment 2 was to determine a miss distance that would best discriminate among the display conditions.

Independent Variables

The three displays tested were NS-NS, HN-HN, and HS-HS. The one display from experiment 1 that was not tested was NN-NN which had neither predictor nor history. The miss distances were 3,000, 6,000 and 9,000 ft.

Experimental Design

Eight pilots were tested, two per day for 4 days. The order of the nine display/miss-distance conditions was randomized for each pilot pair. On all trials the intruder appeared at 60 sec and disappeared at 44 sec to encounter. Errors were averaged over the 12 straight and 12 curved encounter types. These data were analyzed in a 2-encounter-types × 3-displays × 3-distances v 8-pilots ANOVA for a randomized block, repeated measures design.

Results

Table 5 shows the percentage error data for each display and miss distance for both straight and curved encounters. The results of experiment 2 were generally consistent with those of experiment 1. With a 3,000-ft miss distance, numerous errors were made with all displays on both straight and curved encounters. Errors decreased with increasing miss distance, with the largest drop occurring between 3,000 and 6,000 ft. Straight encounters had fewer errors than curved encounters. The straight predictor was only a help on the straight encounters.

TABLE 5

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR THE WINDOW TASK FOR THREE DISPLAYS WITH STRAIGHT AND CURVED ENCOUNTERS. THE MISS DISTANCES ARE 3,000, 6,000 and 9,000 FT. THE DATA IN EACH CELL ARE THE AVERAGES FOR EIGHT PILOTS AND 12 ENCOUNTERS

Display Conditions			Encounter Types					
			Straight			Curved		
History	Predictor	Code	3,000 ft	6,000 ft	9,000 ft	3,000 ft	6,000 ft	9,000 ft
History	None	HN-HN	34 ±4	18 ±3	13 ±5	36 ±4	20 ±4	24 ±4
None	Straight	NS-NS	25 ±6	8 ±3	3 ±2	43 ±2	30 ±5	17 ±4
History	Straight	HS-HS	22 ±7	6 ±3	5 ±2	39 ±3	25 ±4	17 ±3

The ANOVA indicated a significant difference between straight and curved encounters ($F(1,7) = 27.08$, $p < 0.01$) and between miss distances ($F(2,14) = 68.16$, $p < 0.001$). With the variance associated with miss distance eliminated from the variance associated with display, no significant difference was found for display types ($p < 0.1$). However, the interaction between encounter type and display type was significant ($F(2,14)$, $p < 0.025$). All other interactions were nonsignificant.

All eight pilots preferred the display condition showing both history and straight predictor (HS-HS). Four pilots rated the history-only (HN-HN) display as the least liked. The history and straight predictor display was considered easiest to interpret by seven pilots. One pilot commented that the history and straight predictor display made it possible to "see a trend develop." Six pilots thought the history and straight predictor display was the fastest to interpret and three pilots indicated that the display with only history took the most time.

EXPERIMENT 3

The objective of experiment 3 was to investigate the effect of two additional predictive elements on pilot error rate, especially for those encounters involving curved trajectories.

Independent Variables

Three types of predictors were tested: straight (HS-HS), curved (HC-HC), and relative (HN-HR). Condition (HS-HS) has been described previously. The curved vector (HC-HC) was ground referenced and curved proportional to turn rate. It predicted where the aircraft would be in 32 sec if it maintained its current turn rate. In the curved encounters used in these experiments, this predictor always predicted correctly. However, it would be in error if the aircraft turn rate varied during the prediction interval, for example, when an aircraft rolls out of a turn.

The other new display (HN-HR) indicated the position of the intruder aircraft in 32 sec relative to own-ship assuming that both aircraft maintained their current track angles. Mathematically, the vector is derived by subtracting own-ship's straight ground-referenced vector from the intruder's straight ground-referenced vector. The resulting vector is displayed on the intruder symbol (see fig. 1d). If neither aircraft is turning, the relative vector indicates the direction of the intruder across the CDTI. If the extrapolation of this vector lies on own-ship, the intruder is on a collision course with own-ship. If the extrapolation of the vector lies in front (in back) of own-ship, the intruder will pass in front (in back) of own-ship. The interpretation of this display symbol is not as simple if either aircraft is turning. This symbol has been proposed for some collision avoidance displays (ref. 4). All displays included the ground-referenced history. Miss distances were 3,000 and 6,000 ft.

Experiment Design

Four pilots were tested, two per day for two days. The order of the six display/miss-distance conditions was randomized for each pilot pair. Errors were averaged over the 12 straight and 12 curved encounter types. These data were analyzed in a 2-encounter-types × 3-displays × 2-miss-distances × 4 pilots ANOVA with randomized blocks and repeated measures. The curved encounter data were further analyzed in three categories with four encounters each: only own-ship turning, only intruder turning, and both turning.

Results

As can be seen in table 6, fewer errors were made on all displays on the straight encounters. The use of the relative predictor on straight encounters resulted in an error rate that was comparable to the error rates on both the straight and curved ground referenced predictor. The error rate on the curved encounters with curved predictors was much less (6% and 8%) than that for straight or relative predictors (about 33%).

ANOVA indicated a significant difference between encounter types ($F(1,3) = 115.12$, $p < 0.01$) and display type ($F(2,3) = 14.57$, $p < 0.01$). No significant difference was found for miss distance ($p > 0.25$). Interaction between encounter and display type was significant ($F(2,6) = 13.83$, $p < 0.01$). All other interactions were nonsignificant.

TABLE 6

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR THE WINDOW TASK FOR THE THREE DISPLAYS AND STRAIGHT AND CURVED ENCOUNTERS. THE MISS DISTANCES ARE 3,000 AND 6,000 FT. THE DATA IN EACH CELL ARE THE AVERAGES FOR FOUR PILOTS AND 12 ENCOUNTERS

Display Conditions			Encounter types			
Predictor		Code	Straight		Curved	
Own-ship	Othership		3,000 ft	6,000 ft	3,000 ft	6,000 ft
Straight	Straight	HS-HS	8 ±7	0 ±0	33 ±4	29 ±4
None	Relative	HN-HR	6 ±4	2 ±2	33 ±5	37 ±3
Curved	Curved	HC-HC	4 ±2	4 ±2	6 ±4	8 ±5

[illegible] was conducted on the three types of encounters: own-ship turning, intruder turning, and both turning. These data are shown in table 7. In all instances, displays using curved predictors resulted in the lowest error rates (9%, 0%, and 12%). The highest error rate for all displays was for the case in which both aircraft were turning (46%, 50%, and 12%). Analysis of variance of the curved encounters indicated a significant difference between encounter types ($F(2,6) = 7.0$, $p < 0.05$) and display types ($F(2,6) = 27.22$, $p < 0.001$).

TABLE 7

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR THREE DIFFERENT TYPES OF CURVED ENCOUNTERS: OWN-SHIP TURNING, INTRUDER TURNING AND BOTH TURNING. THE DATA ARE AVERAGED OVER THE TWO MISS DISTANCES

Display Conditions					
Predictor		Code	Type Aircraft Turning		
Own-ship	Othership		Own-ship	Intruder	Both
Straight	Straight	HS-HS	37 ±4	9 ±5	46 ±9
None	Relative	HN-HR	31 ±11	25 ±4	50 ±4
Curved	Curved	HC-HC	9 ±3	0 ±0	12 ±4

For this experiment, all four pilots preferred the curved predictor. Suggested improvements for the curved predictor were to flash the symbol of the intruder that was a threat and to allow the pilot to control the length of the predictor. The curved predictor was considered to be the easiest to interpret. On curved encounters, the relative predictor was considered the most difficult symbol to interpret.

EXPERIMENT 4

The objective of experiment 4 was to investigate the usefulness of a CDTI in which own-ship had a curved predictor and the intruder had the simpler straight predictor.

Independent Variables

A practical problem associated with the display of curved predictors is the need for turn-rate information on other aircraft. In an actual system, turn rate could be: (1) predicted based on the past positions of each aircraft; or (2) data linked from each aircraft to the ground from the ground to each CDTI-equipped aircraft. Method (1) can be done on board the CDTI-equipped aircraft but it introduces prediction errors and time delays. Method (2) is inherently more accurate but requires all aircraft to down link additional information — an expensive solution.

The three display conditions were: (1) straight predictors on both aircraft (HS-HS); (2) curved predictors on both aircraft (HC-HC); and (3) a curved predictor on own-ship and a straight predictor on the intruder (HC-HS). All display had 32 sec of ground-referenced history. Miss distances of 3,000 and 6,000 ft were used.

Experimental Design

Eight pilots were tested, two per day for 4 days. The order of the six display/miss-distance conditions was randomized for each pilot pair. Error rate for each display was totaled across the two types of encounter and the two miss distances. These data were initially analyzed in a 2-encounter-type × 3-displays × 8-pilots ANOVA with randomized blocks and repeated measures. The curved encounters were further analyzed by category in the same manner as in experiment 3 in a 3 × 3 × 8 ANOVA.

Results

Table 8 shows that there was again less error with the straight encounters. For curved encounters, curved predictors on both aircraft (HC-HC) yielded the best performance. The analysis of variance showed a significant difference between the straight and curved encounters ($F(1,7) = 73$, $p < 0.001$), display type ($F(2,14) = 27.41$, $p < 0.001$), and the interaction between them ($F(2,14) = 21.40$, $p < 0.001$).

TABLE 8

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR EXPERIMENT 4 FOR THREE DISPLAYS WITH STRAIGHT AND CURVED ENCOUNTERS. THE MISS DISTANCES WERE 3,000 AND 6,000 FT. THE DATA IN EACH CELL IS THE AVERAGE OF 8 PILOTS AND 12 ENCOUNTERS

Display Conditions			Encounter Types			
History	Predictor	Code	Straight		Curved	
			3,000	6,000	3,000	6,000
Straight	Straight	HS-HS	19 ±4	6 ±3	45 ±3	23 ±3
Curved	Straight	HC-HS	15 ±6	4 ±2	34 ±4	22 ±3
Curved	Curved	HC-HC	17 ±5	4 ±2	13 ±3	6 ±2

As was done previously, the three types of curved encounters were further analyzed. These data are shown in table 9. The average error rate for each display (HS-HS, HC-HC and HC-HS) on the curved encounters were 34%, 28%, and 9%, respectively, averaged across miss distances. An analysis of variance

TABLE 9

PERCENTAGE ERROR PLUS AND MINUS ONE STANDARD ERROR FOR THE THREE DIFFERENT TYPES OF CURVED ENCOUNTERS IN EXPERIMENT 4. THE MISS DISTANCES ARE 3,000 AND 6,000 FT. THE DATA IN EACH CELL ARE THE AVERAGE ERRORS FOR 8 PILOTS AND FOUR ENCOUNTERS

Display Conditions			Type Aircraft Turning					
History	Predictor	Code	Own-ship		Intruder		Both	
			3,000	6,000	3,000	6,000	3,000	6,000
Straight	Straight	HS-HS	44 ±2	28 ±5	38 ±0	13 ±8	59 ±8	38 ±4
Curved	Straight	HC-HS	19 ±7	3 ±3	38 ±0	16 ±3	50 ±4	50 ±0
Curved	Curved	HC-HC	13 ±0	13 ±4	13 ±4	3 ±3	16 ±3	6 ±3

indicated a significant difference between encounter types ($F(2,14) = 28.38$, $p < 0.001$), display conditions ($F(2,14) = 37.06$, $p < 0.001$), and the interaction between them ($F(4,28) = 7.22$, $p < 0.001$).

Three post hoc comparisons were conducted on type of curved encounter over each level of the display variable. Results indicated that there was a nonsignificant difference in error rate for type of encounter when both intruder and own-ship had curved predictors. When only own-ship had the curved predictor, there was a significant difference in error rate between the three types of encounter ($F(2,14) = 21.9$, $p < 0.001$), with straight vectors on a turning encounter resulting in the poorer performance. A comparison of encounters when neither aircraft had a curved predictor indicated a significant difference in error rate ($F(2,4) = 7.08$, $p < 0.001$) indicating that the encounters in which own-ship was turning were the most difficult to perceive correctly.

The display with curved predictors on both aircraft was preferred by six of the eight pilots. This display was considered to be the easiest and fastest to interpret by more than half of the pilots. The display with straight predictors on both aircraft was liked least by five pilots.

QUESTIONNAIRE SUMMARY -- EXPERIMENTS 1 TO 4

Ninety-two percent of the pilots thought a CDTI would be of value in their aircraft. Many pilots suggested that the display would be of most value in congested terminal areas. Although 70% of the pilots thought workload would be increased, 63% still felt the display would be necessary; 91% felt a CDTI would improve safety.

Of the suggestions offered for improving the display the most common response (40%) was curved predictors. Seventy-eight percent of the subjects preferred the CDTI located in combination with the weather radar display rather than as a replacement for the HSI.

The pilots were asked to indicate what additional information they would like available. The most common requests were for the ability to change map scale and to read the altitude, heading, and speed of other aircraft. Weather information and terrain symbols were also frequently chosen. The pilots indicated that they desired altitude, heading, and speed of their own and other aircraft to be displayed continuously.

Thirty-eight percent of the pilots wanted collision avoidance instructions such as "TURN RIGHT" or "CLIMB." On the question of announcing with audio alert of the presence of an intruder aircraft, 50% wanted an alert and 50% did not.

Fifty-eight percent thought 1-sec updates of traffic information would be better than the 4-sec updates used in these experiments. Twenty-one percent of the pilots preferred the 4-sec updates and one pilot felt that 8-sec updates would be adequate.

The pilots were asked to rate the experiment itself in a number of areas. Ninety-one percent thought the instructions were clear. The presence of

another pilot in the cockpit was not considered to be a distraction. Seventy-nine percent of the pilots felt the feedback after each trial was helpful in improving their performance.

CONCLUDING REMARKS

A series of experiments was designed to evaluate candidate CDTI symbols in a dynamic but controlled environment. The following general observations are based on the objective and subjective data from the first four experiments in the series and our experience gained in conducting them.

1. This was a difficult perceptual task. Numerous errors were made on all display and task conditions.

2. Curved encounters in which one or both aircraft was turning were much more difficult than straight encounters.

3. Predictive information was more useful than historical information.

4. Both ground-referenced and own-ship-referenced straight predictors reduced errors on straight encounters.

5. Both ground-referenced and own-ship-referenced straight predictors were difficult to use during curved encounters.

6. The turn rate information provided by the curved predictors reduced errors on curved encounters.

7. The best results were obtained when both aircraft had curved predictors.

It is important that any conclusion the reader may draw from these results take into account the following limitations of the research method: (1) The pilots could devote their uninterrupted attention to the task for 16 sec; (2) neither aircraft changed its speed or turn rate during an encounter; (3) although the display was dynamic, the pilots could not interact with the display to ask for more information about the intruder or change the map scale; (4) pilots could not take over manual control of the aircraft; and (5) the passive nature of the task and the large number of trials resulted in a task that quickly became routine. The first two items should lead to fewer errors than would be expected in a real aircraft. The last three items should lead to more errors than in a real aircraft. It is felt, however, that the relative difference between the displays will remain the same as the task is made more realistic.

REFERENCES

1. Technical paper on Cockpit Display of Traffic Information (CDTI). Submitted to the Sub-Committee on Transportation, Aviation and Weather, Committee of Science and Technology, House of Representatives by the Federal Aviation Administration and the National Aeronautics and Space Administration September 20, 1978.

2. Curry, Renwick E., Will the ATSD Precipitate Display-Assisted Collisions? In: The Role of an Airborne Traffic Situation Display in an Evolving ATC Environment, Lincoln Laboratory, MIT, Mass., Report No. PB-215-714, May 1972.

3. Hart, Sandra G., Content of Symbolity and Format of Cockpit Display of Traffic Information: Pilot Opinion. Paper presented at the 15th Annual Conference on Manual Control, March 20-22, 1979, Wright State University, Dayton, Ohio.

4. Litchford, George B., Avoiding Midair Collisions, IEEE Spectrum, September, 1975.

SESSION 6: DESIGN METHODOLOGY AND APPLICATIONS

Moderator: Mr. Ronald O. Anderson, Air Force Flight Dynamics Lab

AN INDUSTRIAL UTILITIES MODEL FOR RESEARCH IN MAN-PROCESS CONTROL

by
B. Dell Campbell
and
Richard S. Shirley

The Foxboro Co, Foxboro, MA 02035

ABSTRACT

This paper describes an industrial utilities model which has been developed to allow laboratory investigation of a man-process control system. It also summarizes a project being done jointly by The Foxboro Company and Delft University (in the Netherlands). The final objective is to determine the influence of high dimensionality and complexity of an industrial process on the operator's performance. For the study, the process is regarded as linear and time-invariant. Currently, an optimal control model of the process operator has been hypothesized, and the industrial utilities model has been implemented.

INTRODUCTION

The industrial utilities model described in this paper will be used for research in man-process control. The following introduction provides a brief, general background of the initial research objectives.

Most industries are in business to make money. But operation of an industrial process is often constrained by government regulations, market demands and physical realities. This usually translates to the question, "how can I make a quality product, not pollute the environment, keep my equipment running, keep my employees productive and happy, and still make money, all at the same time?" In many industries the answer means using more complex processes and process controls. This results in increased responsibility for the process operator, requiring him to monitor a larger, more remote process: some processes have more than four hundred control loops and over a thousand possible alarms. Providing the process operator with a means of reliably monitoring and controlling these loops and alarms is a major goal of any process control company. Consequently, considerable effort is being spent to learn what the operator does and how he does it. One step toward understanding process operation will be made by modelling the process operator.

Useful models have been developed for describing the man in a man-vehicle system. McRuer and Baron, among others, have devised pilot models for use in flight research (references 1,2,3). Typically, the pilot is modelled as a control element in a closed-loop system. Such a system, modified to fit the man-process situation, is shown in figure 1. The man (operator) observes the process via a display and takes control action as needed.

Unfortunately, the model shown in figure 1 is inadequate for process control. As shown in figure 2 and further described by Beaverstock in reference 4, the process operator is really a supervisor; he must learn the process, monitor its behavior, take appropriate control actions, help define standard operating procedures and plan future control strategies. There are several reasons for this. Industrial processes have time constants ranging from minutes to days, much slower dynamics than aircraft; such slow dynamics force the operator into a supervisory role in which control actions are interspersed with long periods of observation, reflection and boredom. Industrial processes tend to have large numbers of control loops compared with an aircraft; the process operator must decide how to divide his time among these loops. Furthermore, operator observations of the process are primarily through remote sensors connected to instruments on the panel, with no direct sensations such as motion or an out-the-window view; the operator must look at the instruments, decide what the signals mean, and then take action. Figure 3 thus indicates that the process operator is more than the simple control element shown in figure 1: he is also a decision maker.

Some of the inadequacies of the operator model shown in figure 1 have been addressed by recent research. Govindaraj and Kok, among others, have modified pilot models to make them more applicable to process control (references 5,6,7). The work of Kok begins with the BBN (Bolt, Beranek and Newman) optimal pilot model (references 2 and 3, and figure 4). The BBN model is a sophisticated version of the situation shown in figure 1. The model shows the man as a control element in a control loop. It also includes disturbances internal to the man (motor noise and observation noise), an internal processing time delay, a predictor/Kalman filter to estimate the system state, an optimum control law to determine control inputs, and neuromuscular dynamics. Kok further builds on this model to adapt it for process control (reference 6). As shown in figure 5, he includes both an internal model of the system dynamics and an internal model of the display as part

of the process operator, now called an observer/controller. Note that figure 5 tacitly assumes a state estimator, functionally similar to the BBN predictor/ Kalman filter. One concern with the model shown in figure 5 (and reference 6) is that the operator's internal model of the process is not completely defined.

Working with The Foxboro Company, Delft University has undertaken a study to determine the influence of high dimensionality and complexity of an industrial process on the operator's internal model of the process. The study by Delft will be based on laboratory experimentation and the work of Kok. Operators will control a simulated industrial process under controlled conditions to provide the data needed for analysis. The industrial process simulation has been developed by The Foxboro Company and is the primary subject of this paper.

THE PROCESS MODEL

A brief summary of an industrial utilities process is shown in figure 6. The process has three major subsystems: boilers, turbines, and compressors. The four boilers use fuel to produce high pressure steam. The three turbines use the high pressure steam to produce electricity, and exhaust intermediate pressure steam and low pressure steam for use in the plant. The four air compressors use intermediate pressure steam to produce compressed air, and also exhaust low pressure steam for use in the plant.

Figure 7 shows the interconnections of the boilers, turbines and air compressors. Even on this simplified diagram there are sixteen operator inputs and four process outputs. Although each control loop is very sluggish compared to man-vehicle systems, the number of control loops is sufficient to keep the operator busy.

Each of the four boilers shown in figure 7 can be turned on or off by the process operator. The operator also decides which of the boilers to use as the "swing boiler" (the swing boiler is used to adapt to small changes in high pressure steam demand). Figure 8 shows the boiler model in more detail. The swing boiler uses a PID (Proportional-Integral-Derivative) controller to maintain a constant pressure (850 psi) on the high pressure steam header. There are four settings on a PID controller under operator control, of which only the set point is shown in figure 8: the other three settings are the gains on the proportional, integral and derivative controls. The time constants for the boilers are of the order of 300

seconds, and there is an integrator in the loop to model the change from steam flow to steam pressure at the high pressure header.

This is a very simplified boiler model: the dynamics are simplified, the process of water flow to the boiler has been omitted, and the dynamics have been linearized. This simplified approach has been taken throughout the industrial utilities model to facilitate research on the process operator. The number of controls, the interactions between the various utilities components, the alarm system, and the long time constants are the aspects of the process of interest to the Foxboro-Delft study of the process operator model.

Not obvious from the figure 7 (for reasons of clarity) is the fact that each of the three turbines in the turbine subsystem is used in a different way. Consequently, each turbine will be discussed in order.

Turbine 1 (see figure 9) is a condensing turbine: it uses high pressure steam to produce electricity, using more of the steam energy in the process than the other turbines. The steam output from the turbine is below the pressure of the low pressure header and so is condensed and returned to the process as water. Turbine 1 is used to control how much electricity must be purchased. The other two turbines also produce electricity, but the amount of electricity they produce varies because they are used to control the intermediate and low pressure steam headers.

Turbine 2 is shown in figure 10. It is a back pressure turbine, using high pressure steam which goes to the low pressure steam header upon exit from the turbine. In addition, a portion of the steam flow is extracted from the middle of turbine 2 to feed the intermediate pressure (300 psi) steam header. The amount of steam flow extracted for the intermediate header is determined by a PID controller setting the constant, K, as shown in figure 10. K ranges from 0 to 1. As K increases, more steam goes to the intermediate pressure header, less steam goes to the low pressure header, and less electricity is generated. The intermediate pressure header is represented as an integrator.

Turbine 3 is the same as turbine 2, except that it is used to regulate the low pressure (40 psi) steam header. As shown in figure 11, the PID controller is reverse-acting. In other words, a portion, K, of the steam flow is extracted

for the intermediate header; if more steam is needed for the low pressure header, then K is decreased. As for turbine 2, K ranges from 0 to 1. The low pressure header is represented as an integrator.

Examination of figures 10 and 11 shows that turbines T2 and T3 tend to work against each other. T2 is used to regulate the intermediate pressure header and T3 is used to regulate the low pressure header. If T2 has more steam extracted for the intermediate pressure header (increased K in figure 10), then less steam goes to the low pressure header. If T3 is used to provide more steam to the low pressure header, then less steam is extracted for the intermediate pressure header (decreased K in figure 11). An unstable situation is avoided by deliberately detuning the low pressure control system, thus allowing the intermediate pressure control system to predominate during transients.

The four air compressors are shown in figure 12. They are driven by intermediate pressure steam and exhaust to the low pressure header. Each compressor can be turned on or off by the process operator as needed to meet compressed air demand. The compressors are controlled by a MOCS (Multiple Output Control System), which essentially sets the combined compressor output to meet total demand and also equalizes the output of each compressor. The operator may directly control one or more of the compressors by placing them in manual mode. The air header is represented by an integrator.

The operator must maintain safe operating conditions. To help him, there are several automatic alarms as summarized in table 1. In case steam pressures get out of balance, there are manually set letdown valves from the high pressure header to the intermediate pressure header, and from the intermediate pressure header to the low pressure header, as well as a vent from the low pressure header to the atmosphere.

The operator must not only run the plant safely, he must also run it efficiently. Specifically, the cost of running the plant is a function of the fuel burned, the purchased power, and the steam vented to atmosphere. In addition, there is a cost penalty for not remaining within acceptable tolerances for the three steam pressures, and the air pressure. For the purchased power, a contract exists: if too little power is used, the entire contracted amount must still be paid for; if too much power is used, the extra power costs twice as much as the contracted

power; and a surcharge is paid based on the highest peak power during the year.

The operator of the industrial utilities model has a large number of controllable parameters which are listed in table 2. In monitoring and controlling the plant, he must observe the process and alarms, decide what to do, and then use one of the controllable parameters to drive the process in the desired direction. The operator's task will be complicated by varying plant demand for electric power, intermediate pressure steam, low pressure steam and compressed air.

From an energy point of view, there are two inputs to the system and four outputs. The inputs are fuel and purchased electricity; the outputs are intermediate pressure steam, low pressure steam, electricity and compressed air. This is, indeed, an energy conversion system. Current technology offers sophisticated energy conservation schemes (see Shinskey, reference 8), but these are well beyond the scope of the simplified model being described.

In conclusion, and with reference to figure 7, the operator's task is to:

- maintain safe operating conditions
- provide a steady supply of high pressure steam to the utilities process
- provide a steady supply of intermediate steam to the plant and the utilities process
- provide a steady supply of low pressure steam to the plant
- provide electric power to the plant
- operate the utilities process efficiently

SUMMARY

An industrial utilities model has been developed for use in laboratory experiments involving man-process control systems. It is sufficiently realistic and complex to pose a challenge to the operator, yet simple enough to be implemented on a minicomputer. The model will be used by Delft University to study the effects of high dimensionality and complexity of an industrial process on the operator's internal model of the process.

TABLE 1

UTILITIES PROCESS ALARMS

Measurement

1. High pressure steam header pressure
2. Intermediate pressure steam header pressure
3. Low pressure steam header pressure
4. Purchased power
5. Fuel to boiler 1
6. Fuel to boiler 2
7. Fuel to boiler 3
8. Fuel to boiler 4
9. Steam to turbine 1
10. Steam to turbine 2
11. Extraction from turbine 2 (intermediate pressure steam)
12. Extraction from turbine 2 (low pressure steam)
13. Steam to turbine 3
14. Extraction from turbine 3 (intermediate pressure steam)
15. Extraction from turbine 3 (low pressure steam)
16. Steam to compressor 1
17. Air from compressor 1
18. Steam to compressor 2
19. Air from compressor 2
20. Steam to compressor 3
21. Air from compressor 3
22. Steam to compressor 4
23. Air from compressor 4
24. Compressed air pressure
25. Vented steam flow
26. Intermediate pressure steam flow
27. Low pressure steam flow
28. Electricity to plant

TABLE 2

OPERATOR-CONTROLLABLE PARAMETERS

(The first 16 parameters are shown on figure 7.)

1. Fuel to boiler 1
2. Fuel to boiler 2
3. Fuel to boiler 3
4. Fuel to boiler 4
5. High to intermediate pressure steam letdown
6. Intermediate to low pressure steam letdown
7. Low pressure steam to atmosphere vent
8. Steam to turbine 1
9. Steam to turbine 2
10. Steam to turbine 3
11. Extraction valve from turbine 2
12. Extraction valve from turbine 3
13. Steam to air compressor 1
14. Steam to air compressor 2
15. Steam to air compressor 3
16. Steam to air compressor 4
17. Choice of swing boiler
18.-21. Parameters of boiler PID controller (including set point)
22.-25. Parameters of turbine 1 PID
26.-29. Parameters of turbine 2 PID
30.-33. Parameters of turbine 3 PID
34.-37. Parameter of air compressor MOCS
38.-41. On/off switches for air compressors
42. Auto/Manual switch for air compressor MOCS
43. Manual control for air compressors

REFERENCES

1. McRuer, D., D. Graham, E. Krendel and W. Reisener, "Human Pilot Dynamics in Compensatory Systems," AFFDL-TR-65-15, (1965).

2. Baron, S., and D.L. Kleinman, "The Human as an Optimal Controller and Information Processor," IEEE Transactions on Man-Machine Systems, Vol. MMS-10, No. 1, pp 9-17, (1969).

3. Baron, S., D.L. Kleinman, D.C. Miller, W.H. Levison and J.I. Elkind, "Application of Optimal Control Theory to the Prediction of Human Performance in a Complex Task," AFFDL-TR-69-81, (1970).

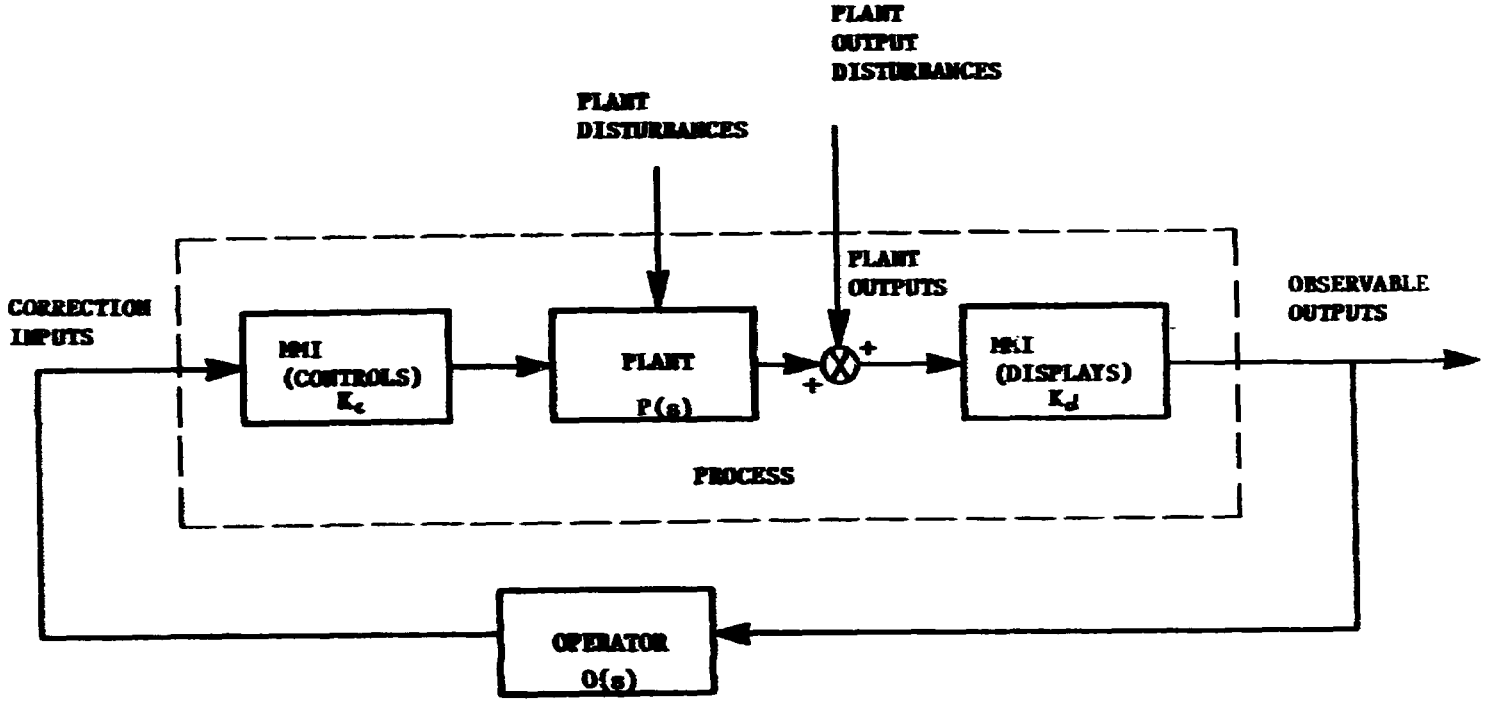
4. Beaverstock, M.C., H.G. Stassen and R.A. Williamson, "Interface Design in the Process Industries," pp 258-265, Proceedings of the Thirteenth Annual NASA-University Conference on Manual Control, held June 15-17, 1977, at the Massachusetts Institute of Technology, Cambridge, MA 02139, (1977).

5. Govindaraj, T., and W.B. Rouse, "Modelling the Human as a Controller of Slowly Responding Systems with Preview in Multitask Situations," Proceedings of the Fourteenth Annual NASA-University Conference on Manual Control, held April 25-27, 1978, at the University of Southern California, (1978).

6. Kok, J.J. and R.A. van Wijk, "Evaluation of Models Describing Human Operator Control of Slowly Responding Complex Systems," doctoral thesis at Delft University, Netherlands, Delft University Press, (1978).

7. Kok, J.J., and R.A. van Wijk, "A Model of the Human Supervisor," pp 210-216, Proceedings of the Thirteenth Annual NASA-University Conference on Manual Control, held June 15-17, 1977, at the Massachusetts Institute of Technology, Cambridge, MA 02139, (1977).

8. Shinskey, F.G., Energy Conservation through Control, Academic Press, Inc., New York, N.Y., (1978).



STRUCTURE OF THE OPERATOR PROCESS SYSTEM

Figure 1

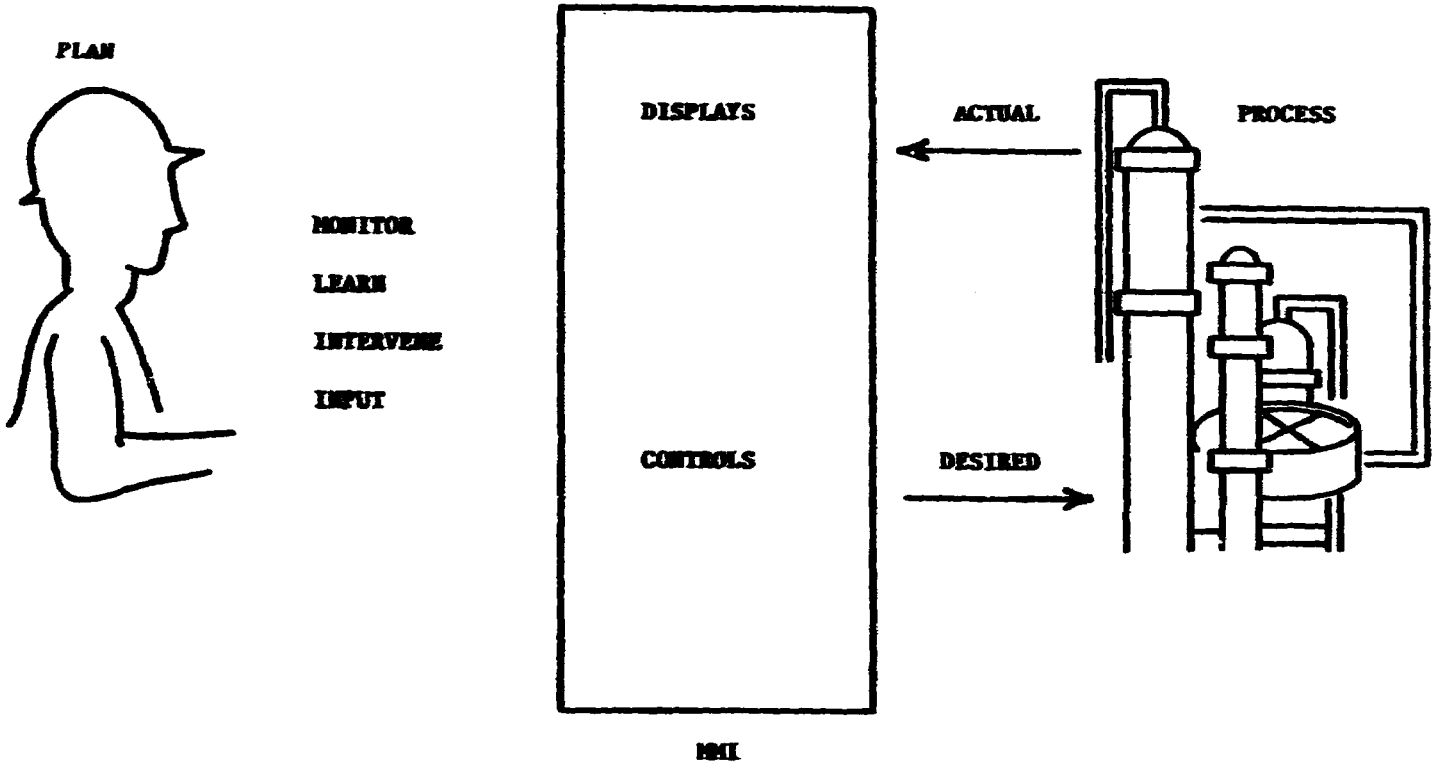
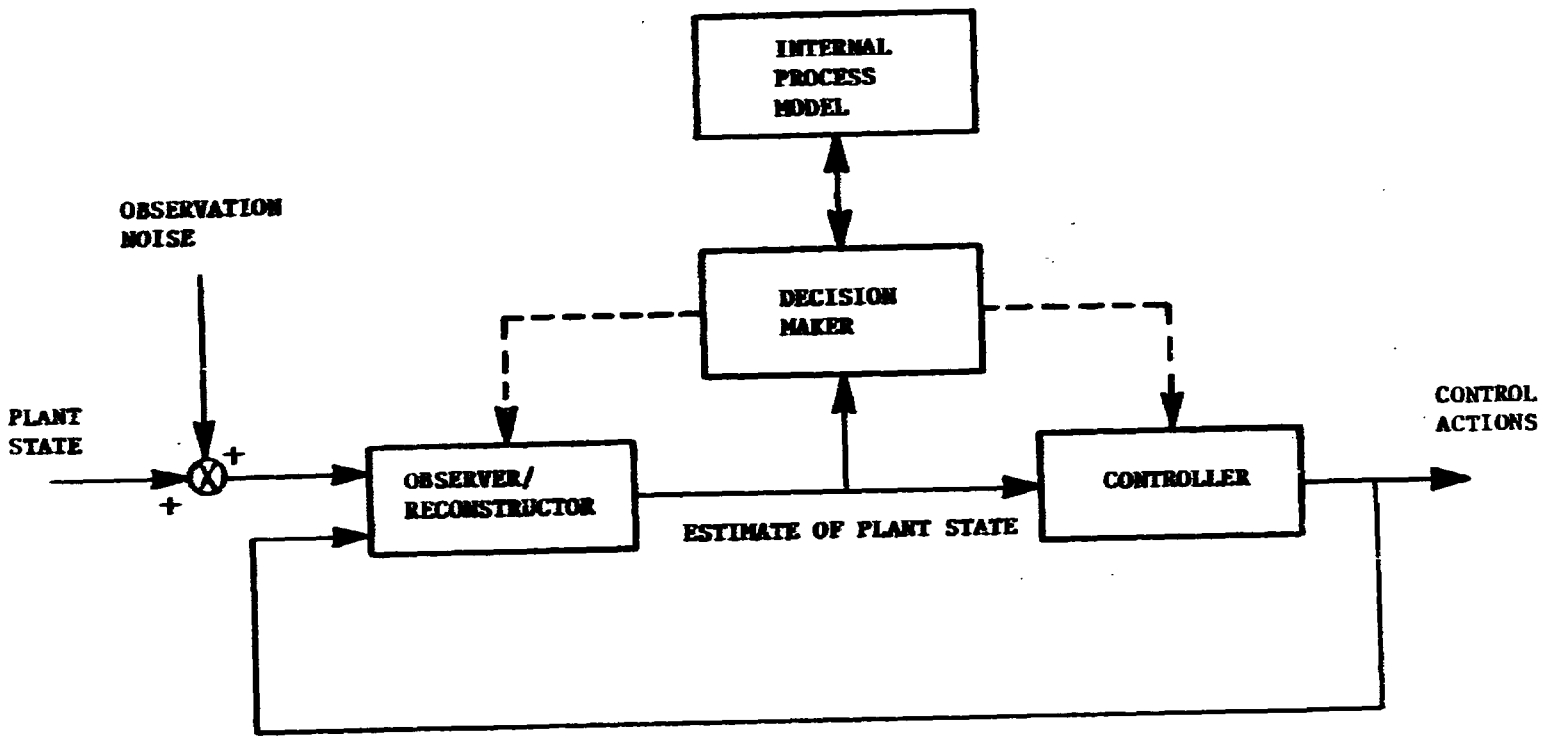


Figure 2

THE TASKS OF THE HUMAN OPERATOR



MODEL OF THE SUPERVISING OPERATOR'S BEHAVIOR

Figure 3

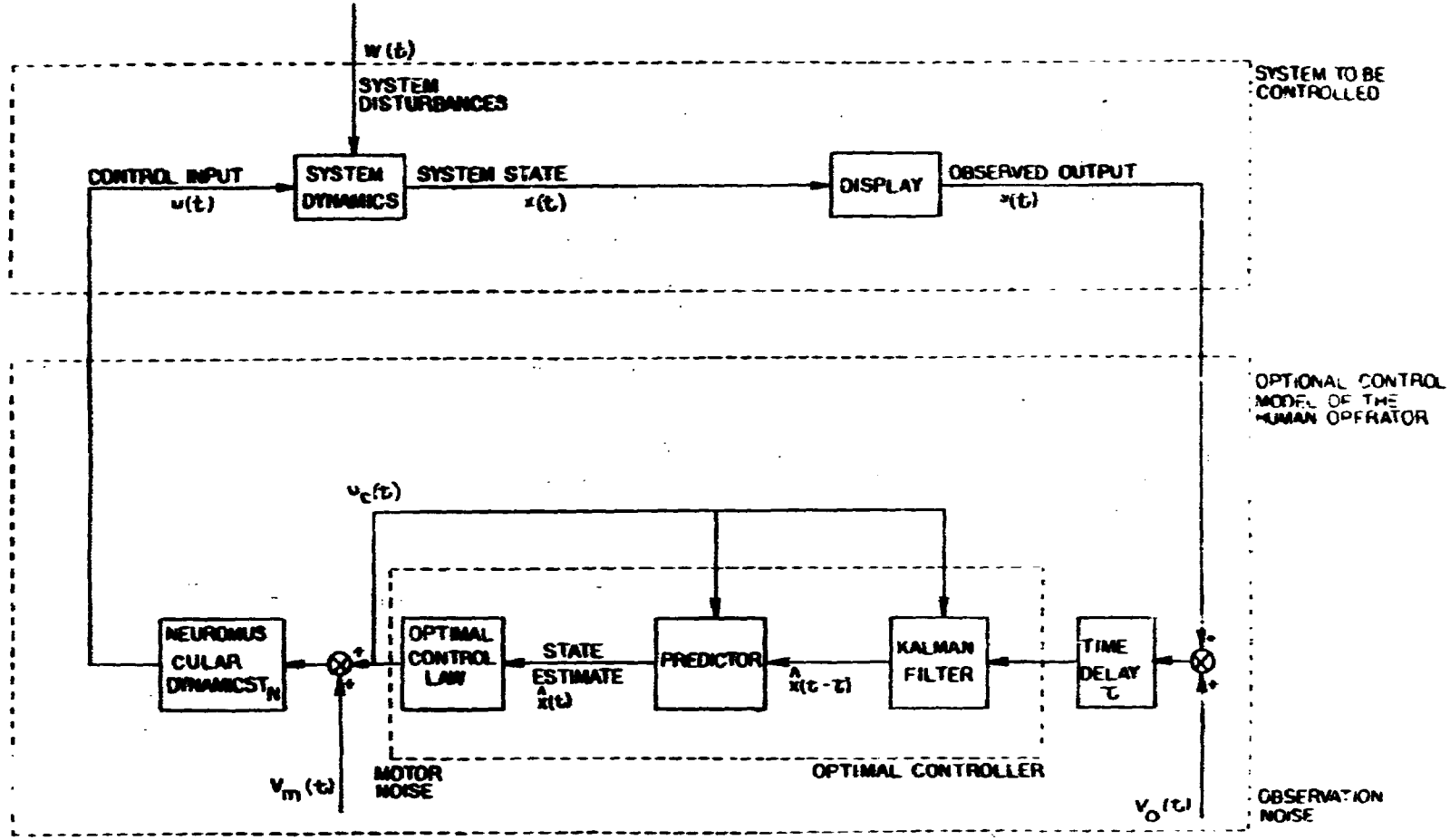


Figure 4

THE BBN OPTIMAL CONTROL MODEL

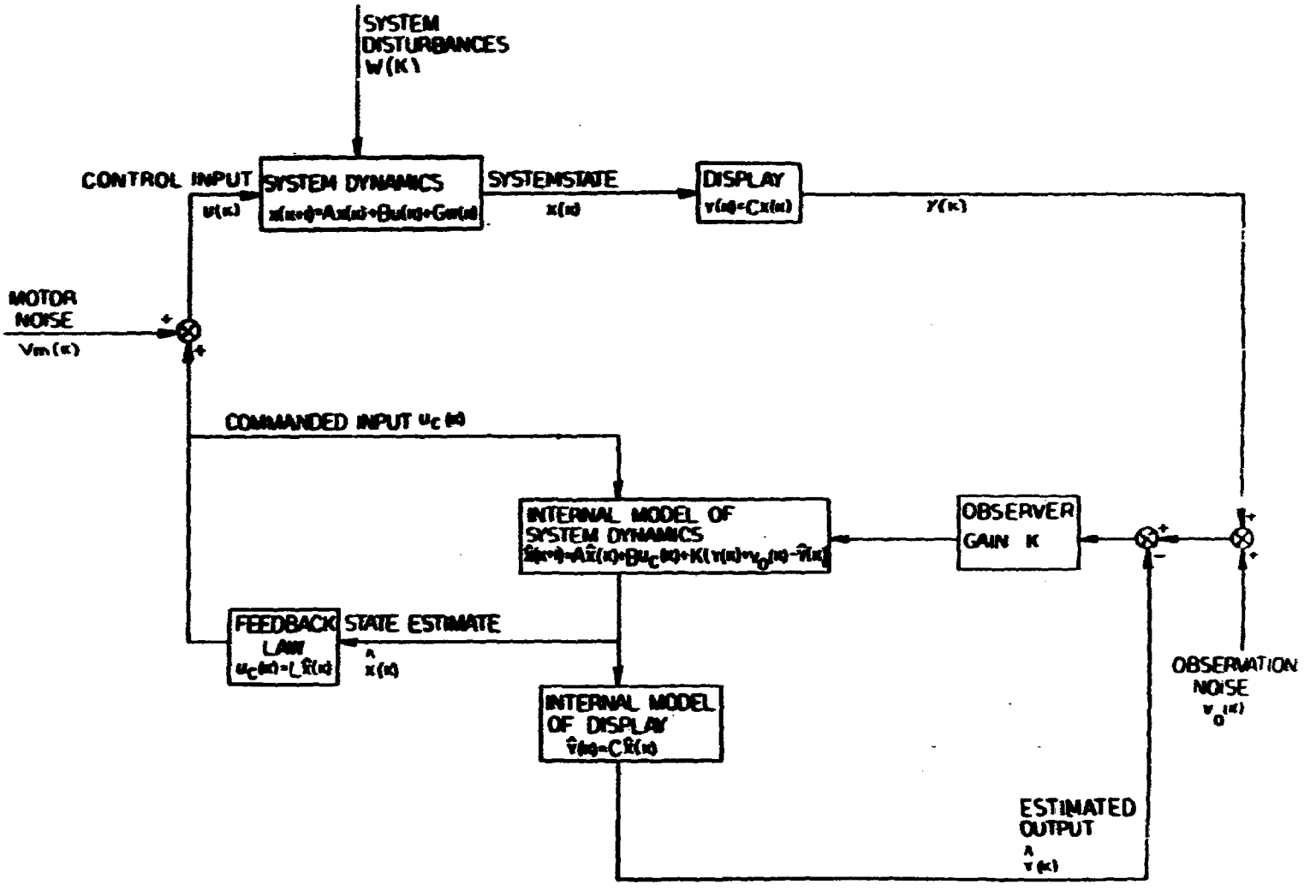
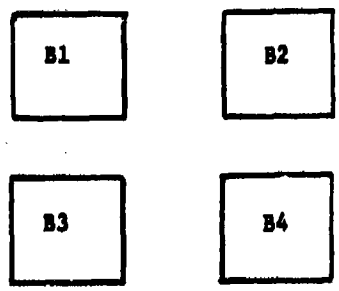


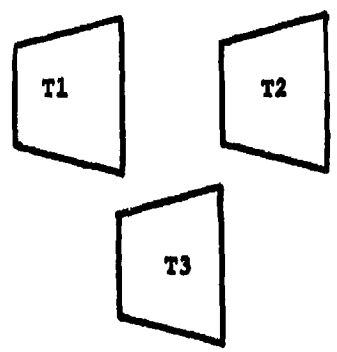
Figure 5

THE OBSERVER/CONTROLLER MODEL

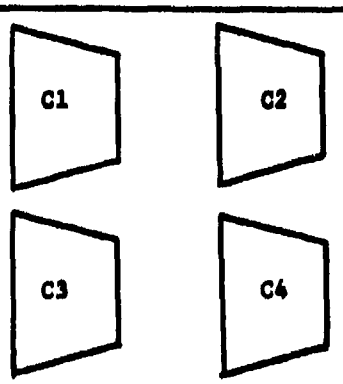
SUMMARY: AN INDUSTRIAL UTILITIES MODEL FOR RESEARCH IN MAN-PROCESS CONTROL



Four boilers -
input: gas fuel
output: high pressure steam



Three turbines -
input: high pressure steam
outputs: electricity
intermediate pressure steam
low pressure steam



Four compressors -
input: intermediate pressure steam
outputs: compressed air
low pressure steam

Figure 6

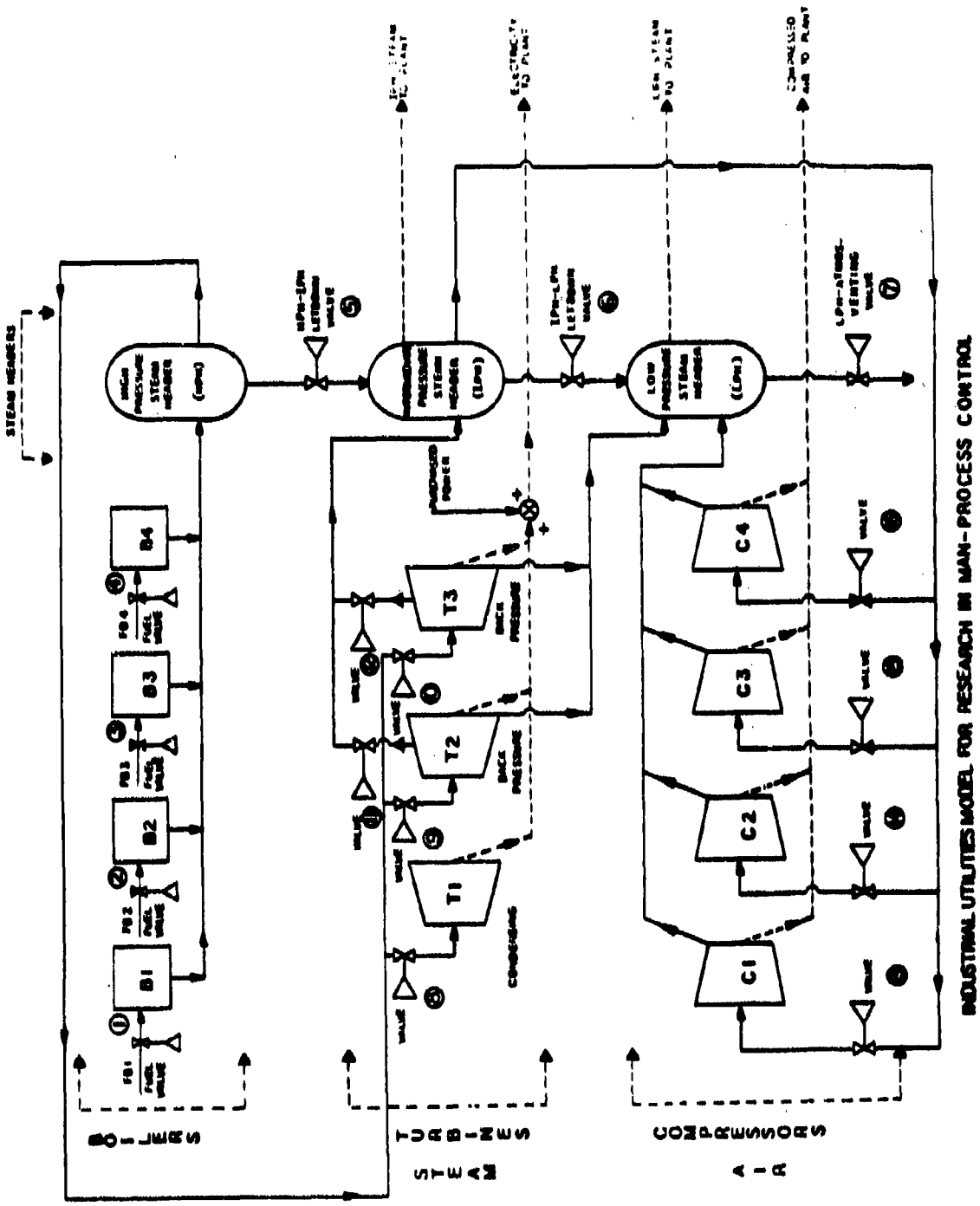
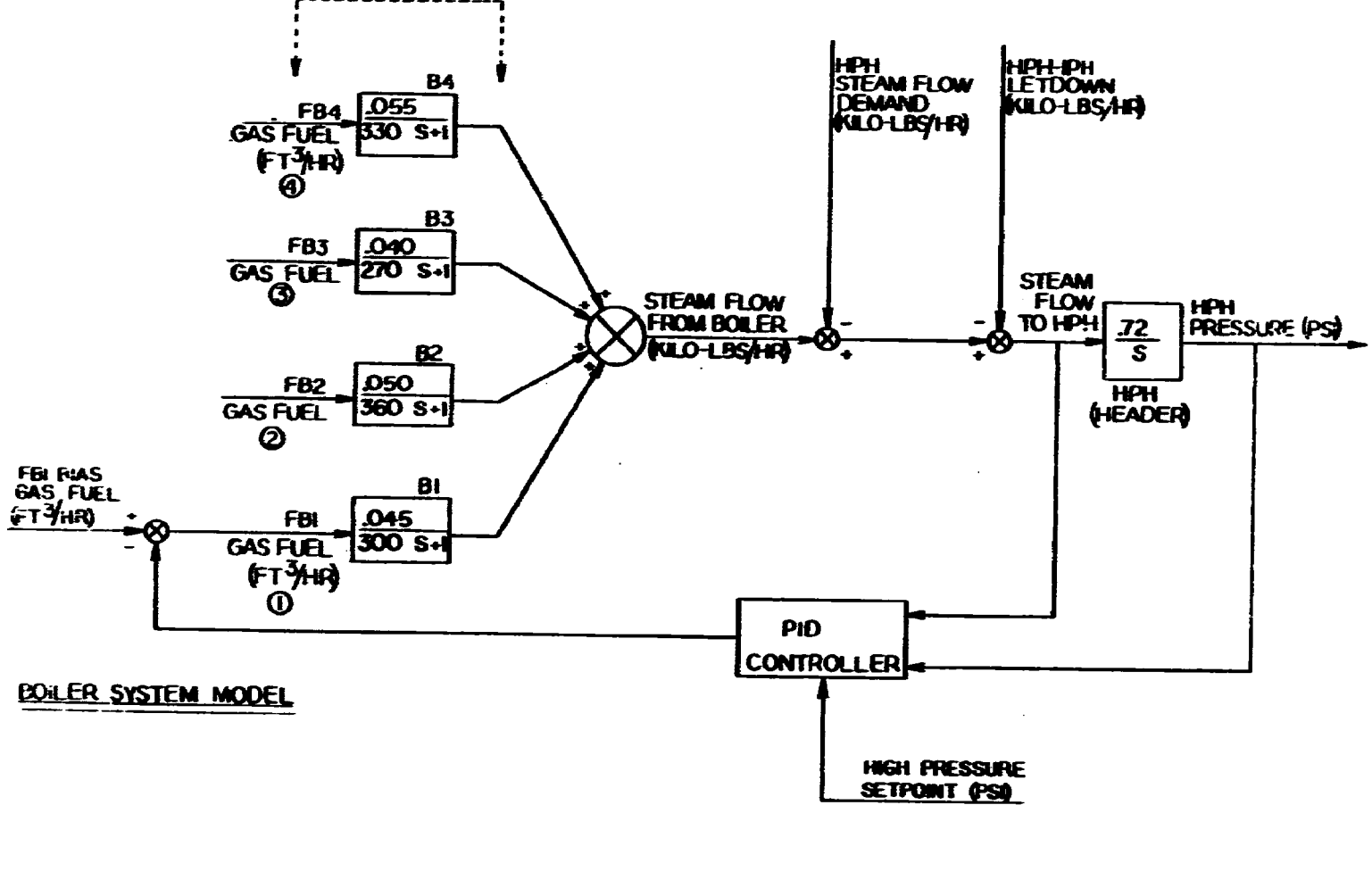
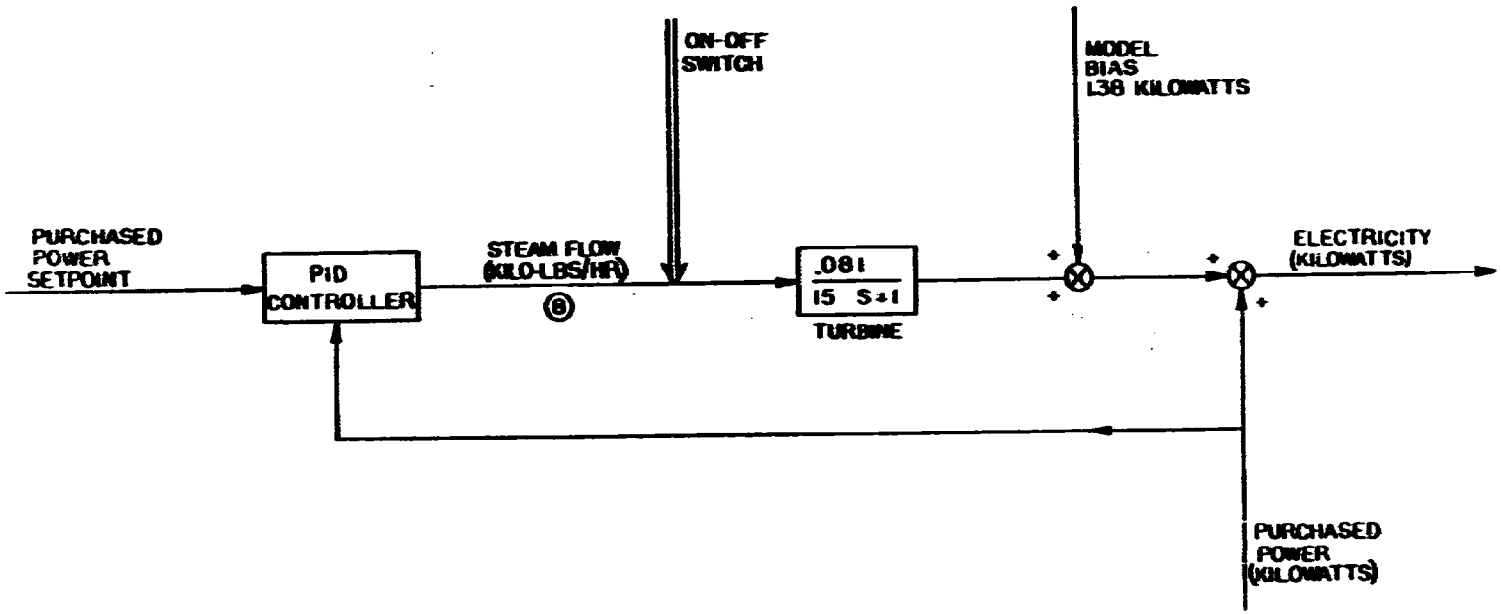

INDUSTRIAL UTILITIES MODEL FOR RESEARCH IN MAN-PROCESS CONTROL
B1
B2
B3
B4
T1
T2
T3
C1
C2
C3
C4
BOILERS
STEAM-TURB-GENS
COMPRESSORS
STEAM
A-R

Figure 7



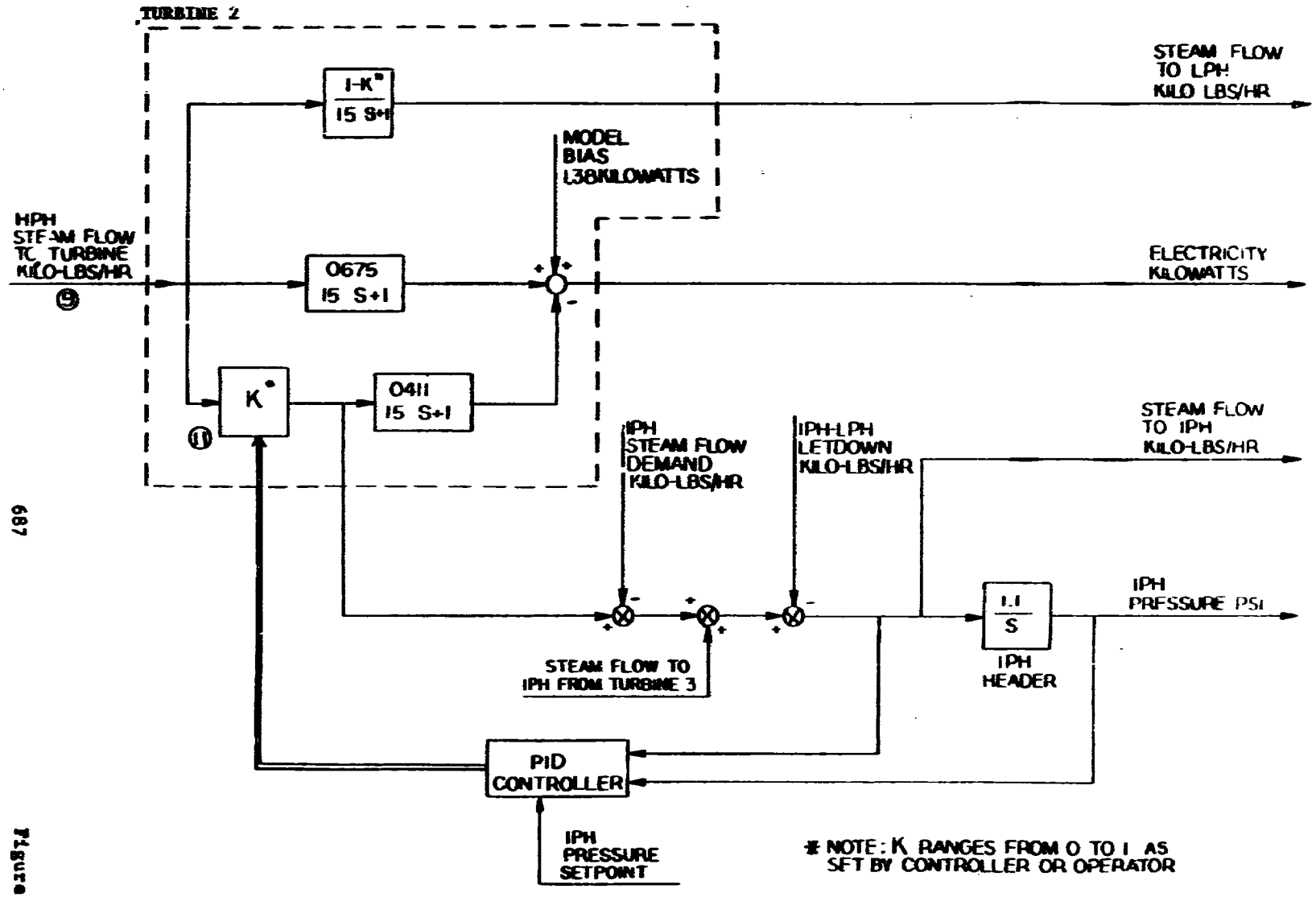
BOILER SYSTEM MODEL

Figure 8



TI CONDENSING TURBINE

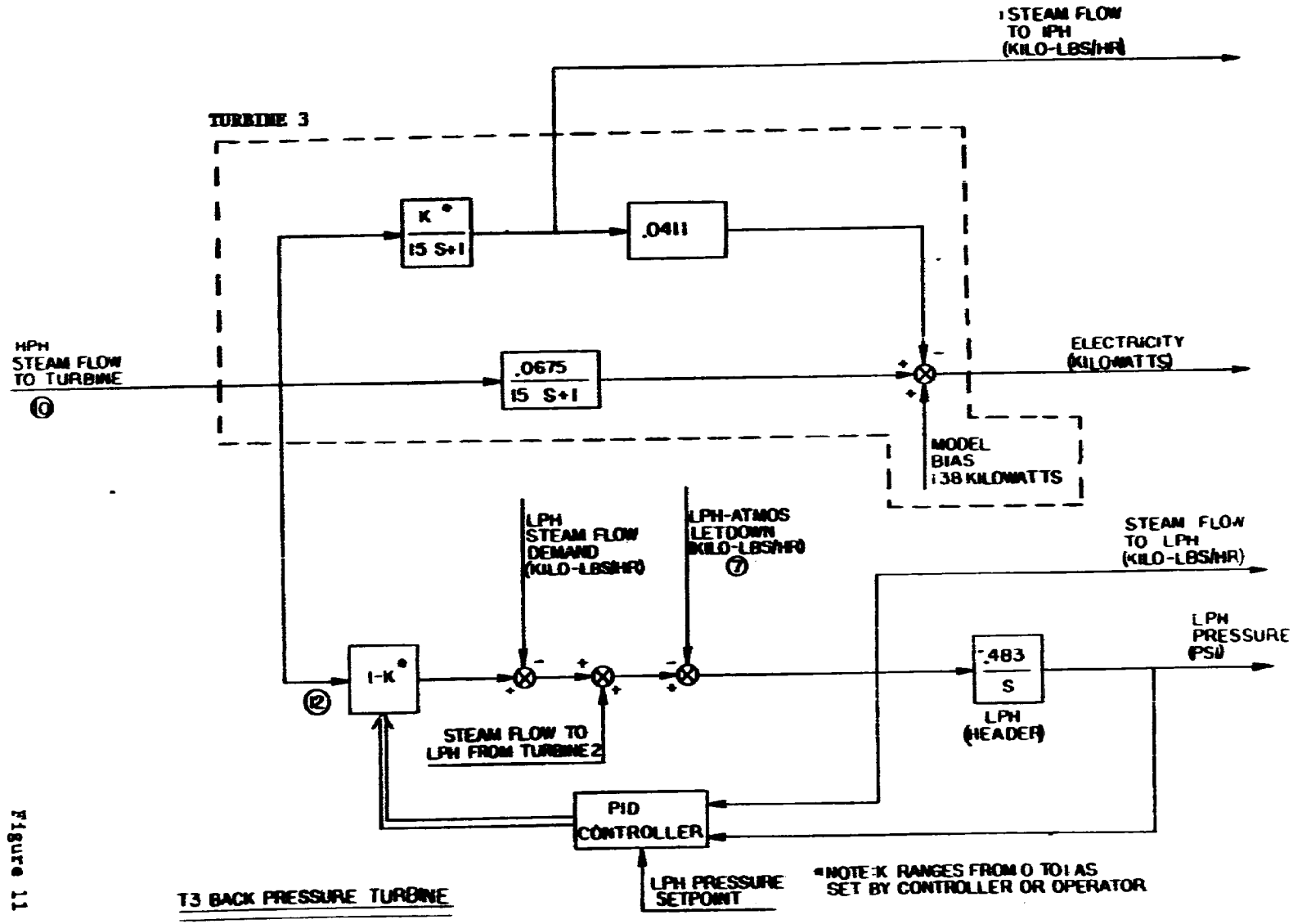
Figure 9



687

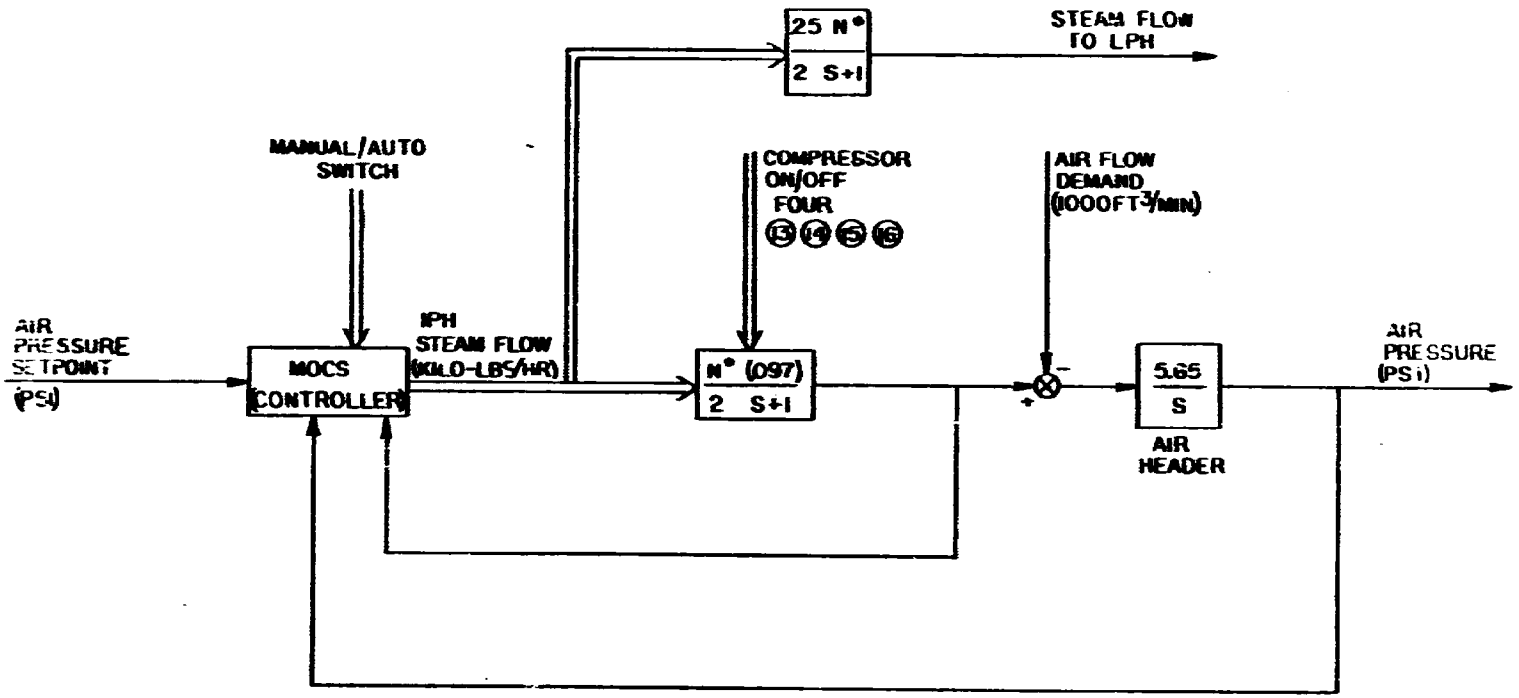
Figure 10

T2 BACK PRESSURE TURBINE



T3 BACK PRESSURE TURBINE

Figure 11



AIR COMPRESSORS

• NOTE: N = # OF COMPRESSORS IN USE, 0-4

Figure 12

The Design and Evaluation of Complex Systems

Dennis B. Beringer

University of Illinois at Urbana-Champaign
Champaign, Illinois 61820

ABSTRACT

The design of complex man-machine systems envisioned for use in evolving environments requires the most rigorous application of a systematic design strategy to insure the best possible interface between system and environment. Careful analysis of operational requirements and component capabilities should lead to efficient allocation of functions and a theoretically optimum synthesis of features. Evaluation of the new configuration in the operational environment, however, may be costly, dangerous, or altogether impractical. Thus an efficient and economical evaluation strategy is needed which will allow system performance to be assessed safely and economically across a wide range of task and environmental variables. Such systematic design and evaluation strategies were applied to an advanced aircraft navigation system using simulation techniques to produce the testing environment. Revised navigation computer/pilot interaction techniques (hardware and procedures) produced shorter training times and better inflight task execution than those obtained with a conventional keyboard entry device, allowing more time to be devoted to manual control of the vehicle. The use of a central-composite design allowed a wide range of environmental variables to be examined using a minimum number of observations.

APPROACHES FOR SUPERVISORY MANAGEMENT OF INFORMATION IN AUTOMATED AIRBORNE SYSTEMS*

By
Randall Steeb
Yee-Yeen Chu
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, California 91367

Abstract

The selection of information for display in airborne systems is a recurrent, subjective decision involving many factors - aircraft state, environmental conditions, operator capabilities, and information acquisition costs, among others. The selection may be characterized essentially as a queueing process with a multi-attribute utility model as a criterion function. Queues of information compete for attention regarding such areas as aircraft dynamics, system faults, threat information, and communications. Continuous build-up in uncertainty regarding each unattended process is maintained from time distributions of event occurrence. The paper describes a method by which the next information to service is selected using a criterion function based on Multi-Attribute Utility (MAU) concepts, which was found useful in ascertaining the value of each information source. Issues regarding queue ordering, service disciplines, and aiding types are discussed.

Introduction

Task analyses of the pilot in advanced aircraft have demonstrated the increasing importance of supervisory control functions. Equipped with wide-band data links, multiple sensors, modular expandable avionics, and on-board computers, the next generation of aircraft are quite likely to be capable of carrying out all phases of flight automatically. However, in view of the projected huge increases in information and resources available to the pilot, more effective and efficient use of information becomes important.

The pilot as the airborne system manager faces a variety of information sources and displays -- such as Master Monitor Display, Integrated Multi-Function Display, etc. These displays may be event driven, functional or procedural. The pilot has the responsibility to monitor the aircraft subsystems as well as supervise the autopilot and to detect possible hardware failures and potential hazards. He must constantly respond to action-evoking events such as: communication of information, change of aircraft configuration, and reduction of 4-D guidance errors. Finally, the pilot is required to respond to unexpected events such as identification and avoidance of threats, change of flight plan, establishment of the backup mode, and declaration of emergencies, etc. The pilot is in a multi-information selection situation.

* This research was supported by the Air Force Office of Scientific Research under Contract No. F44620-76-0094.

Information Value Modeling

The selection, acquisition and processing of information are activities that are involved in virtually every aspect of advanced airborne operations. The operator must, under considerable time stress, weigh the probable usefulness and the costs of a variety of competing forms of information -- mission status, track data, environmental information, aerodynamic functioning, etc. These moment-to-moment judgments must often by based on subjective factors, since the decision is normally too complex and dynamic to be analytically tractable.

Use of Multi-Attribute Utility (MAU) models, pioneered by Raiffa and his collegues (1), (2) and by V. Winterfeldt (3), make the information seeking process goal directed, normative and axiomatic. The value of obtaining information is determined by calculating its impact on the expected utility of the subsequent action and control. The information is assumed to change the probability distribution of the consequence sets and, in turn, to revise the expected values of the alternative actions. In its simplest form, the model is a linear additive rule. The utility of an action is considered to be an aggregate of many possible outcomes, each expressed along a set of attributes:

$$EU(a_k) = \sum_{\text{states}} P(z_h) \sum_{\substack{\text{attributes} \\ i}} U_i(a_k, z_h) \qquad (1)$$

where $EU(a_k)$ is the expected utility of action k, $P(z_h)$ is the probability of state z_h occurring, and $U_i(a_k, z_h)$ is the utility function over the i^{th} attribute associated with state h and action k.

The formulation is the result of several key simplifying assumptions. The decision maker is assumed to be risk neutral, so that he is indifferent between the expectation across a set of uncertain outcomes and the uncertain outcomes themselves. Also, the attributes are assumed to satisfy additive independence, allowing the linear additive form of aggregation. Tests for compliance with these assumptions can be found in V. Winterfeldt (3), or Keeney and Raiffa (2).

The sequence of decision stages can be depicted using a decision tree, as shown in Figure 1. The tree is folded back by associating with each possible message the maximum expected utility of the subsequent actions. This folding back represents graphically the process of EU maximization. The favored information source n is then identified by comparing the expectations taken over all possible messages. Based on this framework model adaptation can be achieved by adding a training algorithm tied to the difference between the model-predicted and actually chosen information as described in Figure 2.

An Experiment

An experimental study was performed to test the effectiveness of the adaptive decision model in information management (Steeb, Davis, Alperovitch and Freedy, (4). Individual subjects were required to pilot a simulated aircraft in a changing, hazardous environment. In doing so, the operators were able to select from a variety of forms of information concerning the multiple threats encountered, and take either evasive or aggressive actions.

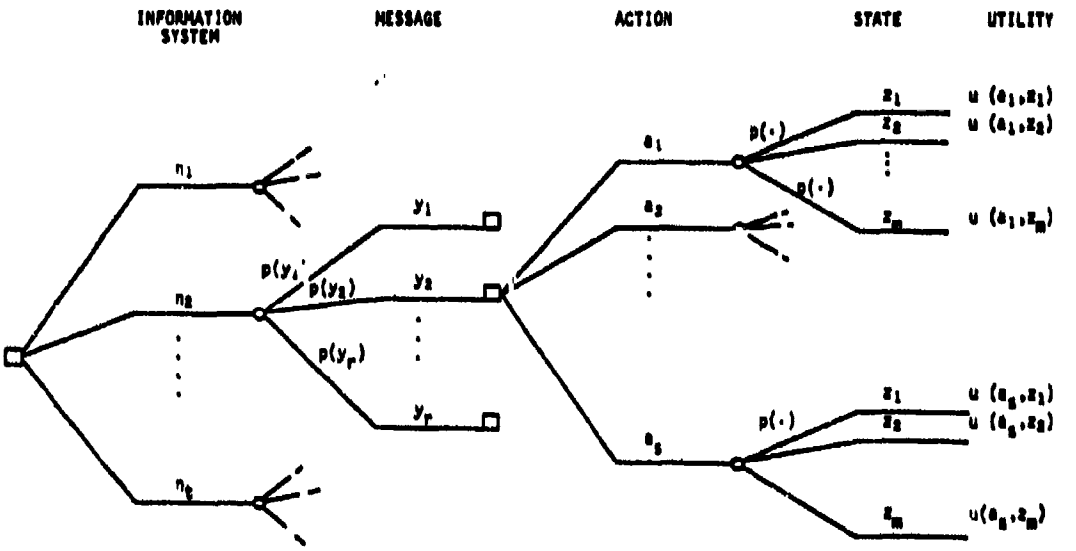


FIGURE 1.
DECISION TREE FOR INFORMATION SELECTION

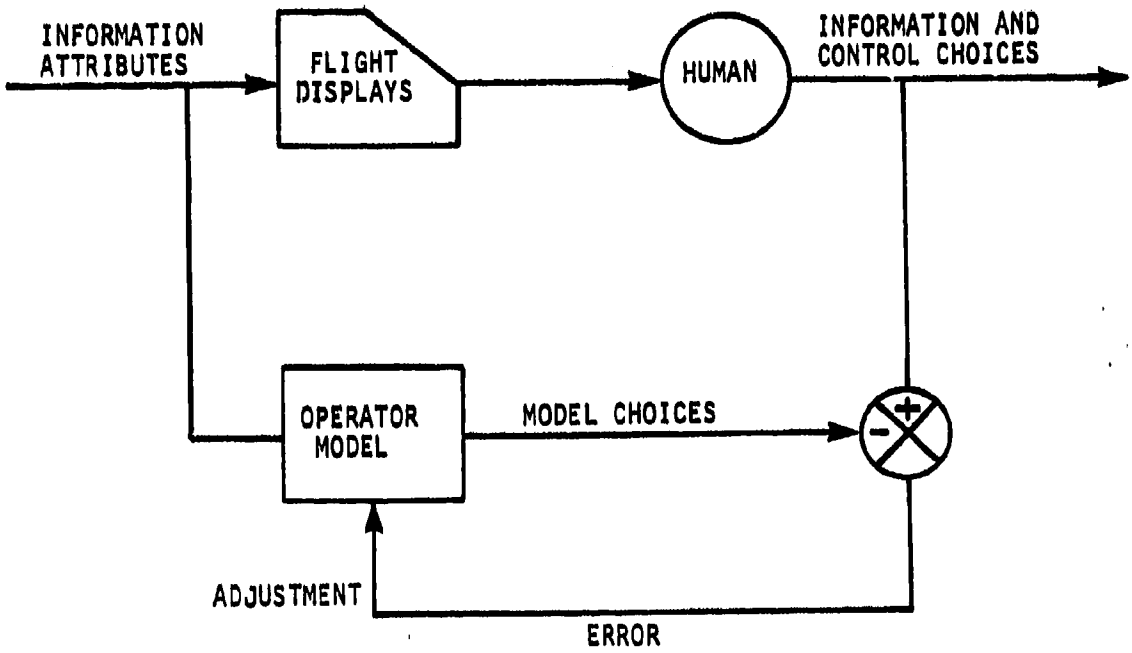


FIGURE 2.
ADAPTIVE STRUCTURE OF DECISION MODEL

The simulation uses a computer-generated CRT display, illustrated in Figure 3. The environment and aircraft are shown as they would be in a moving-map display. Sets of threats appear at random positions at the upper edge of the display and move downward at a constant velocity. The operator can move the vehicle symbol horizontally to one of eleven different pathways to avoid the threats, or he can remain on course and take an aggressive action against one of the threats. The actions open to the operator are primarily decision making in nature. Dynamics of control are minimized since the threat and vehicle velocities are held constant.

The identity and location of the threats can be determined only by exercising one of the five information options. The options differ in their ability to differentiate between and to locate the threats. In those situations where the information option is unable to differentiate between threat types, a combination symbol is displayed.

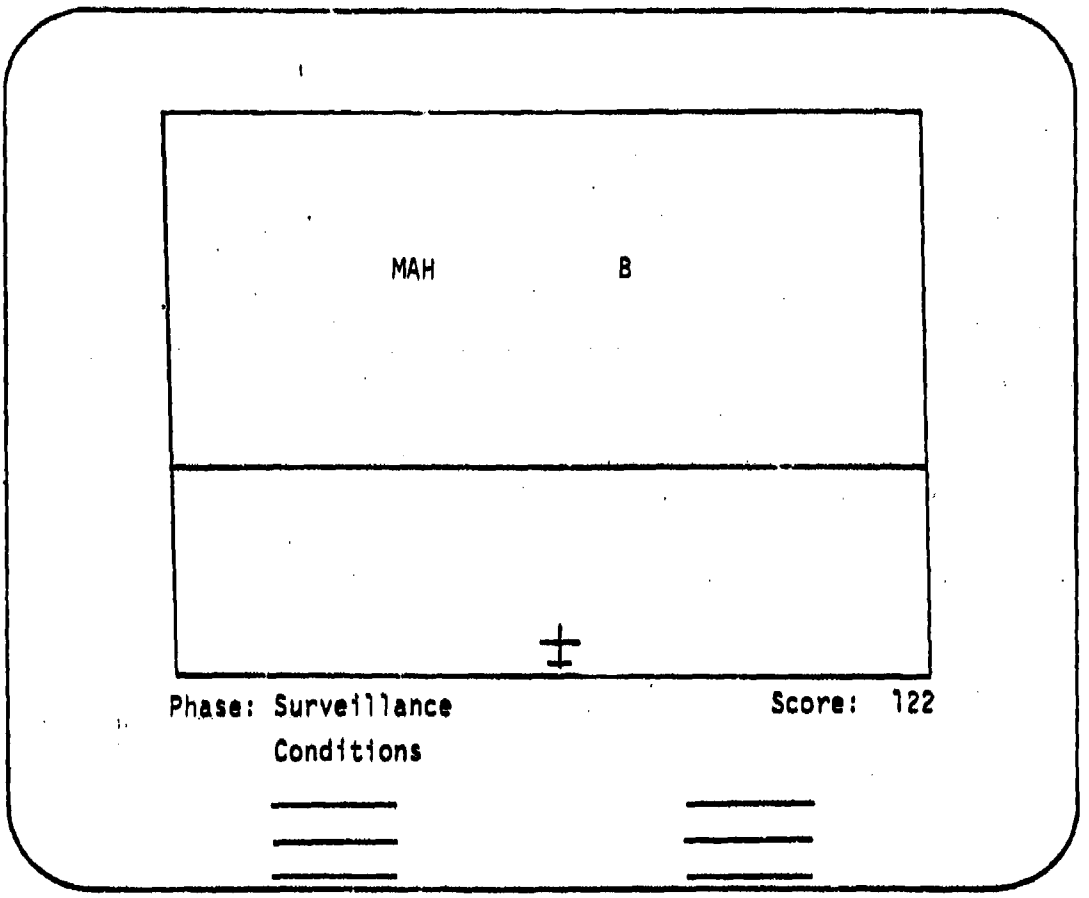
The presentation of conditions is organized into three distinct mission phases -- cruise, surveillance, and aggression. Each phase has set levels of danger, detection, payoffs and information accuracy. Five consequence-related attributes were employed in the decision model: communication cost, time delay for information display, detection, vehicle loss, and offensive gain. Two forms of weight estimation were used -- adaptive estimation from observed behavior and off-line estimation from direct elicitation.

Figure 4 shows the performance score (the payoffs less damages and costs) averaged across the 12 participants, attained under each of the three forms of information management - manual, automated/direct elicitation, and automated/adaptive estimation. The figure also shows the performance scores under the two levels of speed stress (4- or 6-second allowance for information selection). Auto/Adaptive Selection shows a 60% increase in performance over manual selection, and Auto/Direct shows a 29% improvement over manual (all differences significant at $P<.05$). Much of the improvement in aiding appears to be due to the significant decrease in manual performance as the speed-stress increases.

Although the Adaptive model employed in this study was constrained to an application of moderate complexity (multiple criteria, probabilistic consequences, and time-varying behavior) the multi-attribute model was found to provide a useful framework for ascertaining the value of contribution of each information source. The contribution estimated for each information source using the decision model proved to be a good estimator of the actual score attributable to that source. This information value was specific to the individual decision maker and the sequence of task circumstances encountered.

Continuous Monitoring and Control Model

The information model discussed earlier handles the discrete choices present in airborne operations, but does not deal with the many continuous behaviors present in monitoring, tracking, etc. Many of these continuous stochastic processes can be modeled by embedding the multi-attribute decision model in a queueing model. Here the time distributions of processes such as system faults, course errors and energy management losses are known, and a queue of potential messages or sampling options are present. The queueing model provides a descriptive framework to accommodate the build-up in uncertainty regarding a given process. The multi-attribute decision model is then incorporated as a criterion



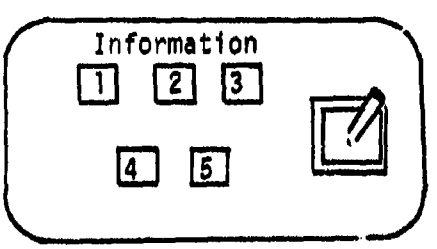


FIGURE 3.
SIMULATED DISPLAY AND COMMUNICATION PANEL

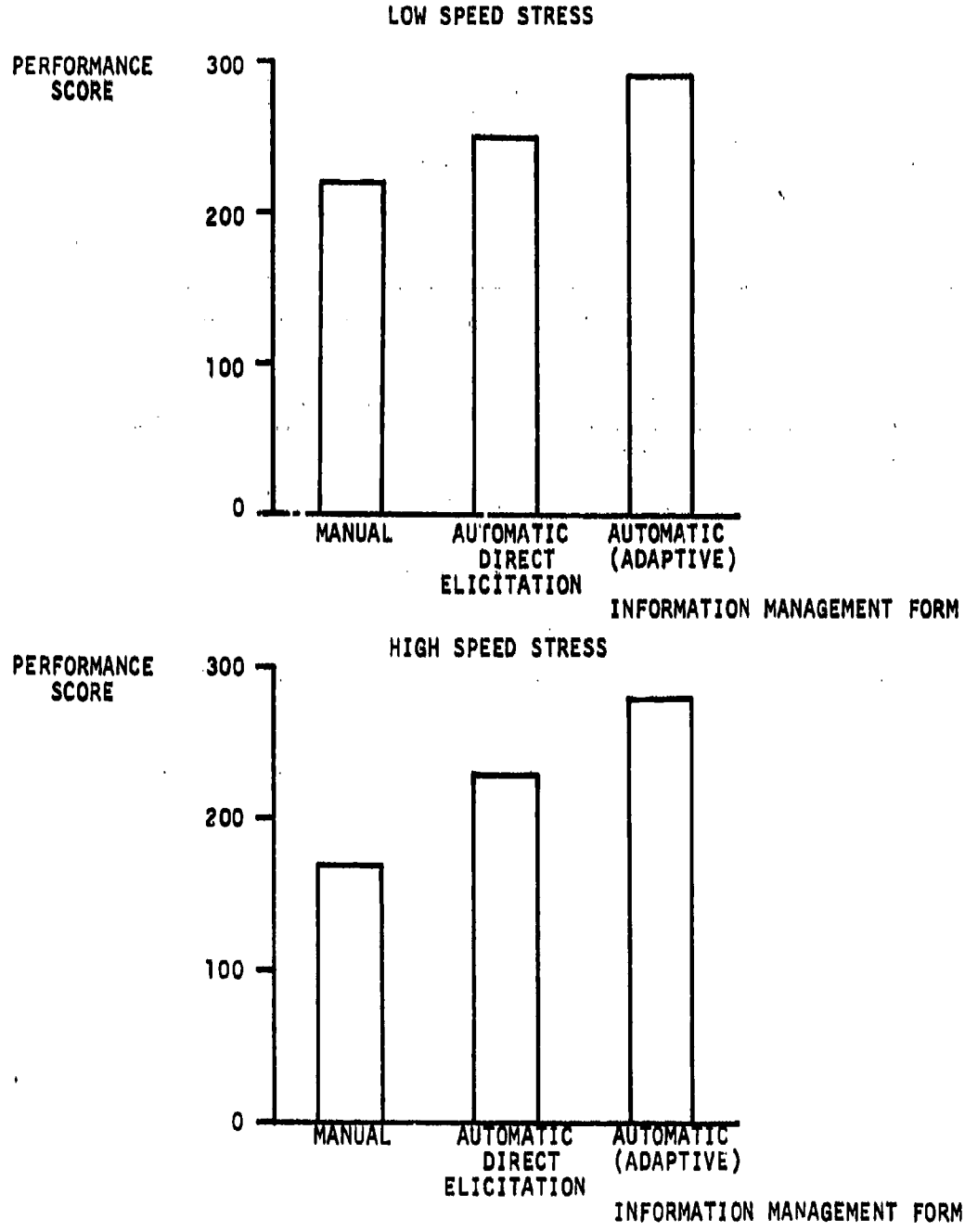


FIGURE 4.
THE RELATIONSHIP BETWEEN INFORMATION SELECTION MODE AND PERFORMANCE SCORE FOR:
(A) LOW SPEED STRESS, AND (B) HIGH SPEED STRESS

function in the queueing model. This section describes the proposed interfacing of the two types of models.

Following previous work by Rouse and Chu (5), a set of assumptions which characterize a basic information selection situation is as follows:

1. The information systems can be represented as K independent processes, with state vectors x_k, k = 1, 2, . . . K.

2. The prior statistics for the observations of system state z_k are given (for example, the joint distributions of the presence of arrivals $p(e_k|z)$, interarrival time distribution $f(t_{e_k}|z)$, and service time distribution $g(t_{s_k}|z)$).

3. The pilot scans the information display (or the pilot is presented with information displays according to some monitoring strategy (queue ordering), the monitoring strategy is based on some set criteria.

4. Action/implementation follows the information selection. False alarms, missed events, and incorrect responses are possible consequences.

This queueing representation of the system features:

1. Time-line analysis of continuous supervisory control.

2. Discrete-time events of multiple processes (monitoring functions).

3. Processes "compete" for attention and "queue" for service.

4. Continuous build-up in uncertainty and risk regarding each unattended process.

5. Multiple operators with independent or overlapping responsibility.

6. Delay time statistics, time-line workload, and system throughput.

7. Integrated, interactive control/display situation.

The framework of the multi-process system assumes that the time distributions of information message arrival and service times are known. With given control of arrival, service or priority discipline (6), the queueing model would predict steady-state system characteristics such as average delay time, throughput, utilization and server occupancy (fraction of time the server is occupied by the message processing). Issues that will be incorporated in the framework include: (1) dynamic (time-varying) priority, (2) task preemption (preemptive-resume), (3) continued sampling, (4) operator load, and (5) uncertainty and other stochastic aspects. The priority discipline for the selection of information arrival will be based on the MAU model. Each of the information arrivals may be characterized by the following pace-related set of parameters (along with other situation-specific characteristics):

1. Frequency of occurrence - arrival rate.

2. Service rate.

3. Cost per unit time delay.

4. Waiting time in message queue.

Also, the value of information must be above some threshold determined by the immediate operator load; i.e., information of low value should not be forced on a highly loaded operator. Some optimum loading level should be maintained. With all these factors taken into consideration, the selection of information is based on the maximization of the MAU associated with each arrival in the queue. The MAU values need only be updated whenever a new message arrives or whenever the environmental phase changes. For automated information selection the messages are selected from the message arrival queue and displayed to the operator. The message pace is determined by the operator's action following the presentation (assuring a full information rate). Therefore, the estimated service time would include information decision and action times.

Figure 5 shows the functional block diagram of the combined MAU and queueing model. The message arrivals are generated from external information sources and transformed in a well-presented format (graphic, schematic, alphanumeric, analog signal, tones, etc.). Each new message arrival causes a reordering of the message queue. The computation of message value is then carried out by the MAU model. The attribute levels and weights may be pre-assigned or estimated adaptively from the previous decision outcomes.

Future Plans

Two types of simulation and testing are planned for the queueing/MAU model. The first simulation will be a purely computational Monte-Carlo simulation. This allows rapid testing of the behavior and performance of the model prior to human subject experimentation. A modification of the current advanced aircraft simulation will then be used to study human operator interaction with various forms of model-based aiding. The simulation of the previous experimental study can be extended to encompass continuous monitoring functions by including additional supervisory tasks. It is planned to supplement the multiple threat task with system fault detection and aerodynamic status tasks. The operator would have uncertainty regarding each monitored process.

This multi-function task, with its use of multiple supervised processes, each with several information options, also lends itself to studies of behavioral issues. Aiding can involve probability aggregation, process recommendation, information recommendation, and automation. The behavioral responses can be observed regarding model confidence, aiding level, operator load, and perceived control.

The forms of display in the task would be those shown in Figure 6. These are simular to that used by Chu and Rouse (7) in their studies of continuous monitoring behavior. Initial work with the task simulation would involve use of automated operators. This would allow rapid checkout of system operation and model function, since the automated operator could operate in fast-time mode. Also, the automated operator could be programmed to exhibit different degrees of inconsistency, time variation, and apparent load. Testing would then proceed to use of human subjects in both single and multiple operator conditions.

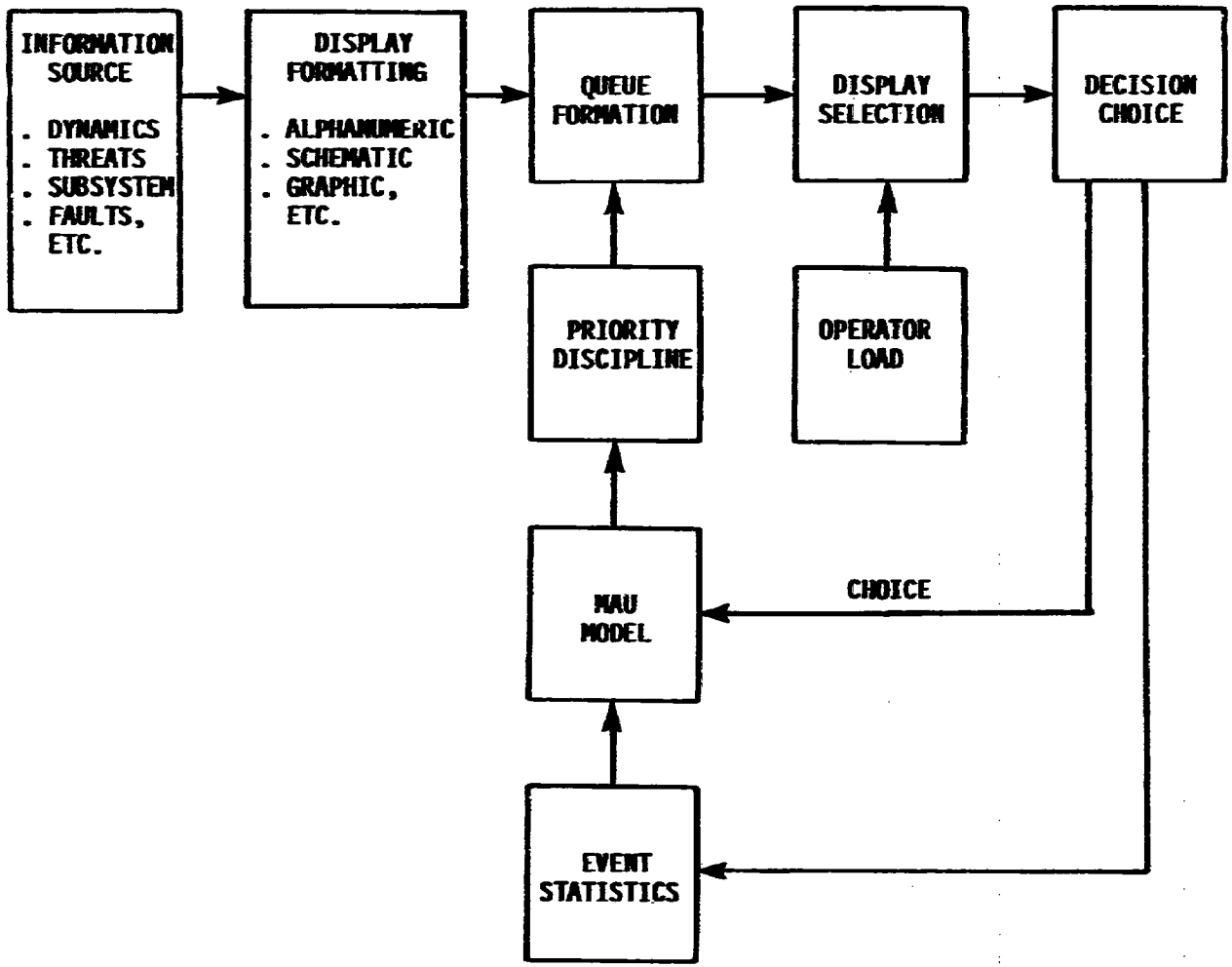
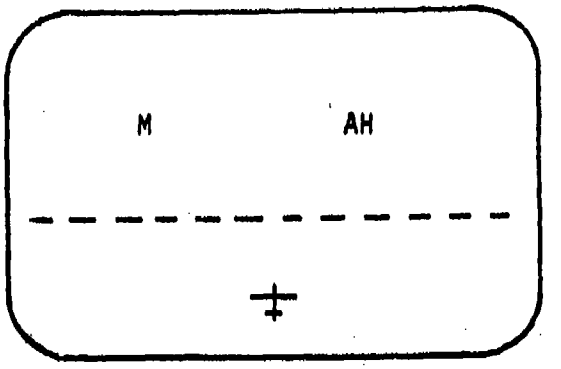
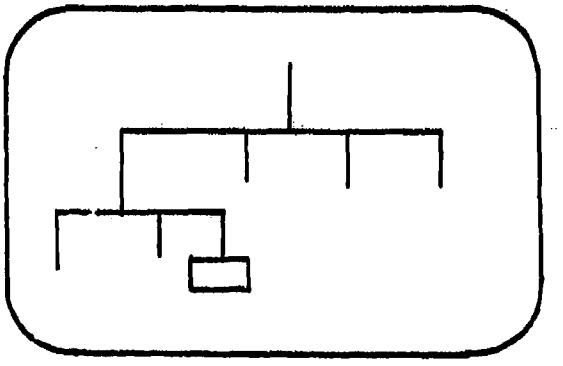


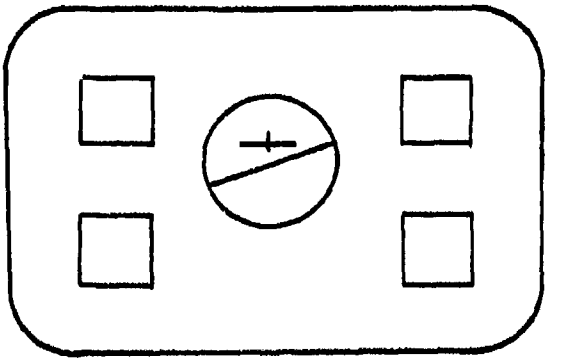
FIGURE 5.
FUNCTIONAL BLOCK DIAGRAM OF QUEUEING/MAU MODEL



(a) Multiple Threat (Full Info Option)



(b) Fault Detection (Graphic Option)



(c) Aircraft Dynamics (Graphic Option)

FIGURE 6.
INFORMATION OPTIONS IN TASK SIMULATION

References

1. Raiffa, H. Rand Corporation (Santa Monica, California). Preferences for Multi-Attributed Alternatives. April, 1969. Memorandum RM-5868-DOR/RC.

2. Keeney, R.L., and Raiffa, H. Decision Analysis With Multiple Conflicting Objectives, Preferences, and Value Trade-Offs. New York: Wiley, 1975.

3. Von Winterfeldt, D. Social Science Research Institute, University of Southern California (Los Angeles, California). An Overview, Integration, and Evaluation of Utility Theory for Decision Analysis. August, 1975. Research Report 75-9.

4. Steeb, R., Davis, K., Alperovitch, K., and Freedy, A. Perceptronics, Inc. (Woodland Hills, California). Adaptive Estimation of Information Values in Continuous Decision Making and Control of Advanced Aircraft. December, 1978. TR-1037-78-12.

5. Chu, Y., and Rouse, W.B. Proceedings of the International Conference on Cybernetics and Society. Optimal Adaptive Allocation of Decision Making Responsibility between Human and Computer in Multi-Task Situations. September, 1977, pp. 168-175.

6. Chu, Y., (ed.) Rouse, W.B. Seminars on Queueing Systems, University of Illinois (Illinois). Optimal Control of Queues. December, 1976. UIL-ENG-77-4002.

7. Chu, Y., and Rouse, W.B. Fourteenth Annual Conference on Manual Control. Pilot Decision Making in a Computer-Aided Flight Management Situation. April, 1978. NASA Conference Publication 2060, pp. 677-690.

EXPERIMENTS ON INTERRUPTED MONITORING OF A SECOND-ORDER PROCESS

by
Everett Palmer
Man-Vehicle Systems Research Division
Ames Research Center, NASA, Moffett Field, California 94035

ABSTRACT

Subjects performed a task in which they could spend only a fraction of their time monitoring a second-order stochastic process for out of tolerance readings. Independent variables were: (1) the process bandwidth (0.2 and 0.4 rad/sec); (2) the fraction of time that could be devoted to monitoring (0.25, 0.50, and 0.75); and (3) the setup cost or time penalty for shifts of attention away from the monitoring task. The subject's monitoring performance and time-sharing strategies are compared to those of a simulation model for the task.

INTRODUCTION

As computers are added to the cockpit, the pilot's role is changing from one of manually flying the aircraft, to one of supervising computers that are doing navigation, guidance, energy management calculations, and automatic flying of the aircraft. This supervisory role adds two tasks to the pilot's duties. One is to monitor the aircraft's performance to make sure the computers are doing their job correctly. The pilot's second task is to insert information and commands into the computers.

The environment in which the pilot interacts with his on-board computer is quite different from other jobs in which an operator interacts with a computer. In a management information system, teleoperator control, or in most human interaction with a computer, the computer is, or can easily be, halted to allow the operator time to think and plan his next input. In those cases, the operator and computer work sequentially. An aircraft that is being controlled in real time by a computer cannot be stopped while the pilot inputs his commands. In that environment, both computer and man must work in parallel. The pilot must interrupt his monitoring to interact with the computer. Conversely, the pilot must interrupt the discrete tasks to monitor. Other characteristics of discrete tasks and monitoring in the cockpit are the following: (1) the discrete tasks are presented at random; (2) they must be accomplished by a certain time — although plenty of time usually is available to do the tasks; (3) attention must be diverted from monitoring for fairly long blocks of time (seconds) to do the discrete tasks; (3) the displays the pilot must monitor show the error between his vehicle's state and the desired state; and (4) these signals are low bandwidth outer loop trajectory variables that must be monitored for out of tolerance conditions.

In this experiment, the essential elements of this task were abstracted to a task in which subjects were presented with a discrete task that had to be

finished by a specified time. Subject to this time constraint, they were asked to divide their attention between monitoring a continuous second-order stochastic process and performing the discrete task so that they maximized the amount of time they were observing the map display when it was out of tolerance.

The objective of this research was to determine how parameters of the process and the discrete task affect both monitoring performance and the pilot's time-sharing strategy. Subjects' performance and strategy were compared to that of a simulation model of a human performing the task.

SIMULATION FACILITY AND TASKS

Monitoring Task

The subject monitored the output of a second-order process driven by white noise. The signal appeared as lateral errors on the simplified area navigation map display shown in figure 1. The subject was told that the errors were due to navigation system noise. The display was defined as being out of tolerance if it had an error greater than "one dot." The process bandwidth determined how predictable the error signal was. The ratio of the tolerance to the standard deviation of the error determined how frequently the error signal was out of tolerance. For these experiments this ratio was set

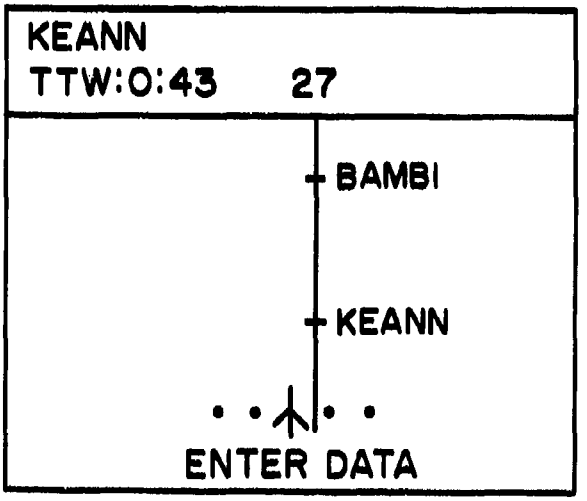


Figure 1. A drawing of the simplified area navigation display used in the monitoring task. The display was generated on a caligraphic computer graphics system. The next waypoint, KEANN, was 43 sec away. Of that 43 sec, 27 sec had to be spent on the discrete task.

to 1.5 and the error signal was out of tolerance about 13.4% of the time. The map display was refreshed 60 times a second and updated 10 times a second. Its motion appeared to be continuous.

During an experimental run, the "aircraft" flew along a route defined by 10 waypoints. The time to the next waypoint (TTW) was displayed in the upper left corner of the map. There were 150 sec from the start of a run to the first waypoint and 75 sec between each subsequent waypoint. When the "aircraft" was 60 sec from the next waypoint, the message "ENTER DATA" appeared at the bottom of the display. The subject, who then had 60 sec to complete the discrete task, switched between the monitoring task and the discrete task by depressing the space bar on a keyboard.

Discrete Task

When the discrete task page was selected, the route line, aircraft symbol, and deviation dots were removed from the display. The two-digit number in the upper center of the display counted down whenever the discrete task page was selected. (This number was also displayed on the map page.) The subject's task was simply to observe that page long enough so that the counter reached zero by the time the aircraft was at the next waypoint. A second experiment, not reported here, investigated a more realistic discrete task in which the subjects had to key in random digits when they were not monitoring the display.

This discrete task had the advantage of allowing precise control of the setup cost of the discrete task and the fraction of time that the subject could monitor the map display. The initial value of the counter determined the fraction of time the subject could monitor the map display. Setup cost is nonproductive time that occurs each time a discrete task is started, either initially or after being interrupted. For example, in entering data with a keyboard, some time is lost when the operator shifts his attention to the keyboard and positions his hands. Setup cost was simulated by adding a delay before the discrete task counter began counting down each time the subject switched to the discrete task page.

Instructions and Performance Feedback

The subject was instructed that his highest priority was to finish the discrete task by the time the aircraft reached the next waypoint. Subject to this constraint, he was further instructed to maximize the amount of time that he was observing the map display when the displayed error was out of tolerance. When the "aircraft" reached the waypoint, the subject was given a score in terms of hits and misses for that waypoint segment. The hit (miss) score was the number of seconds that the error was out of tolerance and observed (not observed) by the subject. This constraint formulation seems to be a more accurate description of real tasks than a task in which the operator is "rewarded" for observing out-of-tolerance signals and doing discrete tasks.

THEORY

In reference 1, three models, based on Smallwood's (ref. 2) internal model concept, were developed for a simpler task in which a subject monitored

the quantized output of a discrete time first-order process. The third model in reference 1 formulated the subject's task as a two-state, finite horizon, dynamic programming problem (refs. 1, 3). The states were the absolute error and the amount of time that had to be spent on discrete tasks before the time horizon. By adding a third state for error rate and quantizing the states, that same formulation could in principle be used to derive normative strategies and performance for the current task. Unfortunately, this dynamic programming approach was computationally impractical.

It was observed in reference 1 that performance and strategies of the optimal dynamic programming formulation were almost identical to those that could be derived from a much simpler "myopic" model. In that model, a decision was made to either continue monitoring or to switch to discrete tasks, depending on which activity maximized the immediate expected reward. The subject received a reward of one unit for observing the map display when it was out of tolerance and a smaller reward of R units for doing discrete tasks for one unit of time. As R was increased, the model devoted a larger fraction of time to discrete tasks. When the performance and strategy of this model were plotted as a function of fraction of time on discrete tasks instead of R, the performance and strategy of the model and the dynamic programming model were essentially identical.

In the remainder of this section, the myopic model is extended to monitoring a continuous second-order process with an infinite horizon. The effect of several process parameters, discrete task parameters, and operator parameters on monitoring performance is illustrated. Finally the model is modified for the specific task performed by the subjects in this experiment in which a specified amount of time had to be spent on discrete tasks by a future time deadline.

In the model, it is assumed that the operator has an internal model of the process and environmental disturbances. He knows the parameters of the process and can use this knowledge to predict the probability that the process will be out of tolerance at a future time given that he knows the displays current error and error rate. For a second-order process with bandwidth ω, the distribution of the displayed error in t seconds is a Gaussian distribution with

$$\text{mean} \qquad m(t) = [e(1 - \omega t) + \dot{e}t]e^{-\omega t}$$

$$\text{variance} \qquad v(t) = 1 - (1 + 2\omega t + 2\omega^2 t^2)e^{-2\omega t}$$

The probability that the error will be out of tolerance at a future time can be calculated from this distribution:

$$P_{out}(e,\dot{e},t,Tol) = \int_{-Tol}^{+Tol} N[m(t),v(t)]de$$

The average probability of being out of tolerance is computed for future times of 1,2,3,..., sec:

$$\overline{P_{out}(e,\dot{e},t,Tol)} = \frac{1}{t}\sum_{i=1}^{t} P_{out}(e,\dot{e},t,Tol)$$

The maximum value of t (the discrete task dwell time T) is found for which $\overline{P_{out}(e,\dot{e},t,Tol)} < R$, the reward for doing discrete tasks. The model then switched to the discrete task for T seconds. If T = 0, the model continues to monitor. Figure 2 shows a plot of the models strategy as a function of display error and error rate for R = 0.032. A simulation of this strategy showed that the model observed 85% of the out-of-tolerance errors while spending 25% of the time monitoring. The average dwell time on the discrete task was 6.3 sec.

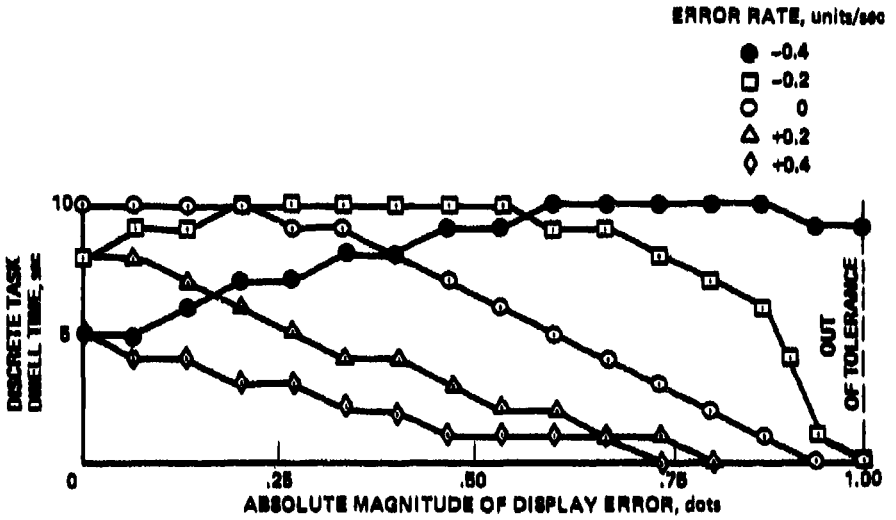


Figure 2. The model's strategy was a function of display error with error rate as a parameter for a discrete task reward of R = 0.032. The process bandwidth was 0.2 rad/sec and the error was out of tolerance 13.4% of the time.

The monitoring performance and other characteristics of this strategy were determined by computer simulation. A series of random numbers was digitally filtered to generate the output of the process. The model sampled this error signal and, based on the above strategy, decided to either keep monitoring or to switch to discrete tasks for T seconds. After dwelling on the discrete task for T seconds the model switched back to the monitoring task. The model always observed the displayed error for a minimum dwell time, usually 1 sec, before making a new decision to either continue monitoring or switch to the discrete task. Simulation runs were made with different values of discrete task reward R. The length of each run was 10,000 sec. The monitoring performance of the model is plotted in figures 3 and 4 as a function of the fraction of time spent on discrete tasks for different process bandwidths and display dwell times. Also shown on those graphs are the performance that would result with perfect information on the displays future error and the performance with no error information. As the bandwidth decreases and the

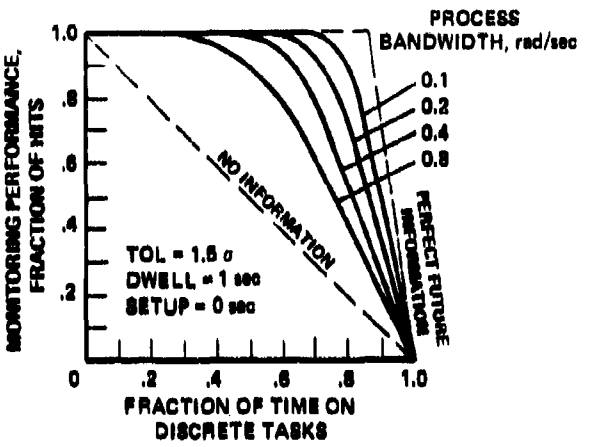


Figure 3. The effect of a process bandwidth on monitoring performance for a second-order process. The error was out of tolerance 13.4% of the time. The minimum display dwell time was 1.0 sec and the discrete task setup cost was 0.0 sec.

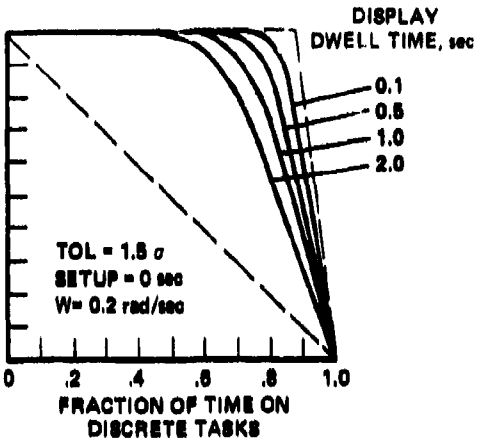


Figure 4. The effect of display dwell time on monitoring performance for a second-order process.

error signal becomes more predictable, monitoring performance approaches that possible with perfect future information. Also, shorter minimum display dwell times result in performance closer to that possible with perfect future information. Further simulation runs showed that if the model always assumed that the display error rate was zero, only a very small decrement in performance occurred. Monitoring performance was also insensitive to varying the models estimate of the process bandwidth from 0.05 to 0.8 rad/sec.

Figure 5 shows the detrimental effect of a discrete task setup cost on monitoring performance. To investigate the effect of setup cost, the simulation model had to be modified as follows. For each switch to the discrete task, the setup cost C was subtracted from the models discrete task dwell time decisions to find the "productive" dwell time spent on the discrete task. Also a lower bound, TMIN, was specified for the minimum discrete task dwell time. TMIN was always at least C + 1. For a given setup cost and reward the value of TMIN was found which maximized performance.

In order to do the finite horizon task presented to the subjects, the above model was modified by adding a feedback loop which adjusted the discrete tasks reward B as a function of the fraction of the remaining time that must be spent on the discrete task. As noted above, as R increases, the model spends an increasing amount of the time on the discrete tasks. The power function, $R = 13.4\ F^{A}$, was empirically found to give a good fit to the relation between R and F (the fraction of the remaining time which must be spent on discrete tasks). The parameter A was set to give the correct value of R for F = 0.75. The values of A ranged from 4.97 to 1.34 for the four bandwidth and setup cost conditions. A test was also performed on the models decisions to make sure that the model did not spend more time on the discrete tasks than was necessary to finish them.

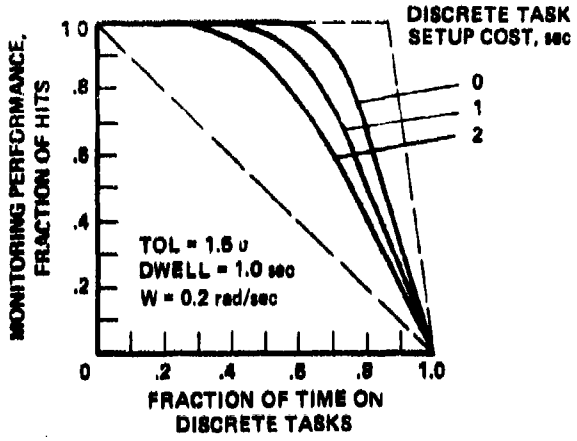


Figure 5. The effect of discrete task setup cost on monitoring performance for a second-order process.

Before the experiment was conducted, this model was run on the identical random sequences that were observed by the subjects in the experiment for each of the experimental conditions. The internal model's bandwidth and order were identical to those of the actual process. The model's perceptual limitations, as described by error and error-rate quantization and observation noise, were set to zero. The minimum dwell time on the map display was set to 1 sec. With those parameters the model should perform better on the average than the subjects in the experiment, unless the subjects use a shorter dwell time on the map display. Monitoring performance and other statistics of the models' strategy, and experimental data from the subjects are shown in figures 6-11. Note that model performance deteriorates when there is a specified time limit. That is to be expected because the model is constrained to spend a specific number of seconds on discrete tasks (15, 30, or 45 sec) out of each 60-sec period. With no time limit, discrete tasks can be postponed indefinitely while waiting for an opportune time to switch to discrete tasks.

EXPERIMENT

The objective of this experiment was to determine how the predictability of the monitored process, the fraction of time that must be devoted to discrete tasks, and the setup cost of the discrete task affected human monitoring performance and attention sharing strategy.

Independent Variables

The key parameter affecting predictability is process bandwidth. The two bandwidths studied were 0.2 and 0.4 rad/sec. These values are within the range of bandwidths of outer loop guidance and control displays in an aircraft cockpit. Three levels of fraction of time required to be spent on discrete tasks were used: 0.25, 0.50, and 0.75. Two levels of discrete task setup cost were used: 0 and 1 sec.

Constant Variables

The following variables were left constant in this experiment:

Variable	Value	Units
Number of displays	1	
Order of process	second	
Damping ratio	1.0	
Error mean	0.0	dots
Error standard deviation	0.666	dots
Number of waypoints	10	
Time between waypoints	75	sec
Time allowed for the discrete task	60	sec
Minimum chunk size	0.1	sec
Page switch time	0.1	sec
Prob error out of tolerance	0.134	

Subjects

Eight male college students aged 18 to 30 were paid for their participation in this experiment. None had previously participated as a subject in any similar experiments.

Experimental Design

A mixed — within subject and between subject — factorial design was used. Fraction of tasks and process bandwidth were chosen as within subject variables because pilots encounter a variety of values of fraction of tasks and also monitor displays with different bandwidths. Setup cost was chosen as a between-subject variable because (1) in a real system, setup cost would be relatively constant for a given system; (2) it was expected that subject strategy would be dependent on setup cost and making it a between subject variable eliminated uncontrolled transfer effect; and (3) it seemed desirable to limit the time any one subject was in the experiment to minimize boredom problems. The two replications per subject were also a within-subjects effect. Four subjects were randomly assigned to each of the two setup-cost conditions. The order in which each subject received the six within-subject variables in each replication was randomized. Each run in a replication was 13.75 min long and consisted of 10 waypoint segments, each 75 sec long, plus a 75-sec beginning segment during which there was no discrete task. At the beginning of each 75-sec segment, the random number generator was reinitialized to 1 of 10 different seeds. For each run a different random permutation of these 10 seeds was used. This procedure made it doubtful that the subjects could memorize the random process while still using the identical process for all conditions.

Procedure

The testing for each subject took place over 3 days. Subjects were run in groups of two, with one subject resting while the other performed the experimental task. On the first day, the general objectives of the experiment

and its relation to aircraft piloting tasks were explained. The subjects were then individually instructed in the operation of the equipment and allowed to monitor each of the bandwidths. At the beginning of each run, they were instructed to attempt to observe the display whenever the error was out of tolerance with the constraint of completing all discrete tasks on time. After passing each waypoint, a display on the CRT showed how well they had done in terms of hits and misses. Any questions on the operation of the equipment were cleared up at that time and subjects were given a rest. They then did each of the six experimental tasks, with a rest between each of those runs. On the second and third days, subjects performed data runs on each of the six tasks. At the end of each run, subjects completed a brief questionnaire.

Dependent Variables

The following data were taken on the 10 waypoint segments of each run: (1) the fraction of hits — the number of seconds the error was observed to be out of tolerance divided by the total number of seconds the error was out; (2) the number of times the subject switched to the discrete task; (3) the mean and standard deviation of the dwell time on the discrete task; (4) a histogram of the dwell times on the monitoring display; (5) the average and standard deviation of the amount of discrete tasks remaining at 45, 30, 15, and 0 sec to the next waypoint; (6) ratings on the questionnaire.

In addition, the process output and the times of each switch between tasks was recorded to allow detailed decision by decision comparisons of subject and model strategies.

RESULTS

Figure 6 shows the fraction of hits for the eight subjects and the model.* Monitoring performance went down sharply when only 25% of the time was spent monitoring. Performance also decreased as the process bandwidth increased and as the error signal became less predictable. The setup cost, or penalty for switching tasks, caused a decrement in performance at all levels of fraction of time on the discrete task. The model's performance is generally equal to or slightly better than that of the subjects. The main exception to this statement is the superior performance of subjects A, B, and C when they could spend only 25% of their time monitoring. As pointed out above, the model was constrained to monitor the display for at least 1 sec. Figure 7 shows the average monitoring dwell time for each subject. These same three subjects (A, B, and C) had a dwell time that was about half that used by the model. The model was run again with a minimum dwell time of 0.5 sec. Model performance improved in the zero setup cost conditions to be equal to or better than that of all the subjects.

*Each data point plotted for a subject in figures 6 to 13 is the average of 20 points for data from the 10 waypoint segments in each run and the two replications of each condition. If a subject did not finish the discrete task in time, his last decision to switch to the discrete task was modified during the data analysis so that the discrete task was always finished on time.

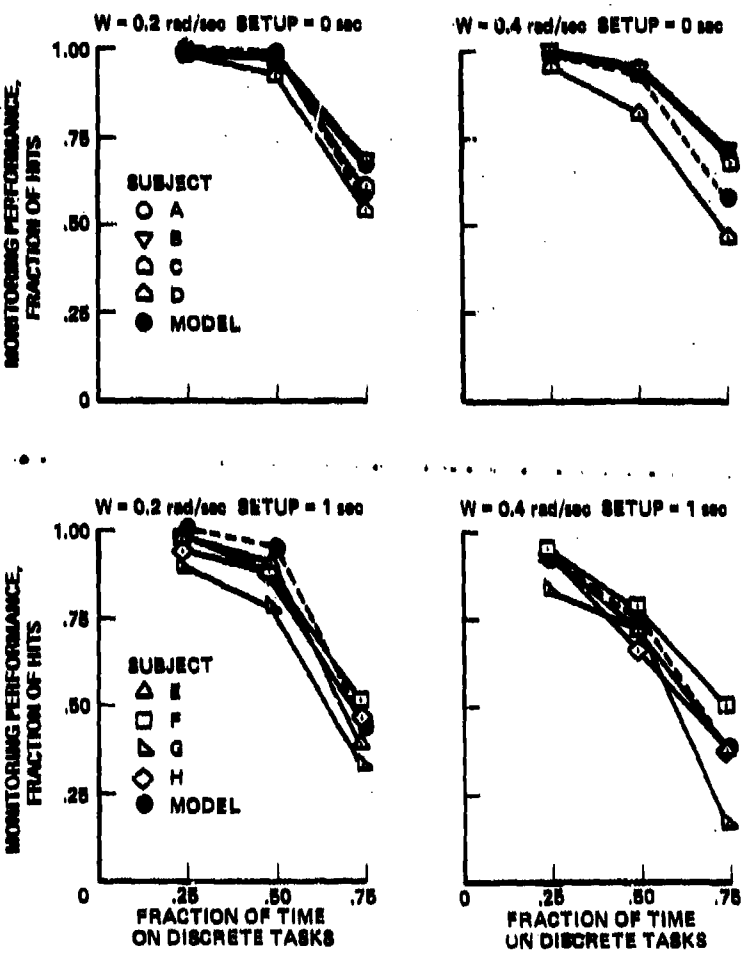


Figure 6. Monitoring performance for the model and eight subjects. Subjects A, B, C, and D performed the tasks with no setup cost and subjects E, F, G, and H performed the tasks with a 1-sec setup cost. Fraction of hits is the amount of time the error signal was observed out of tolerance divided by the total amount of time the error was out of tolerance. The heavy broken line is the model prediction for each condition.

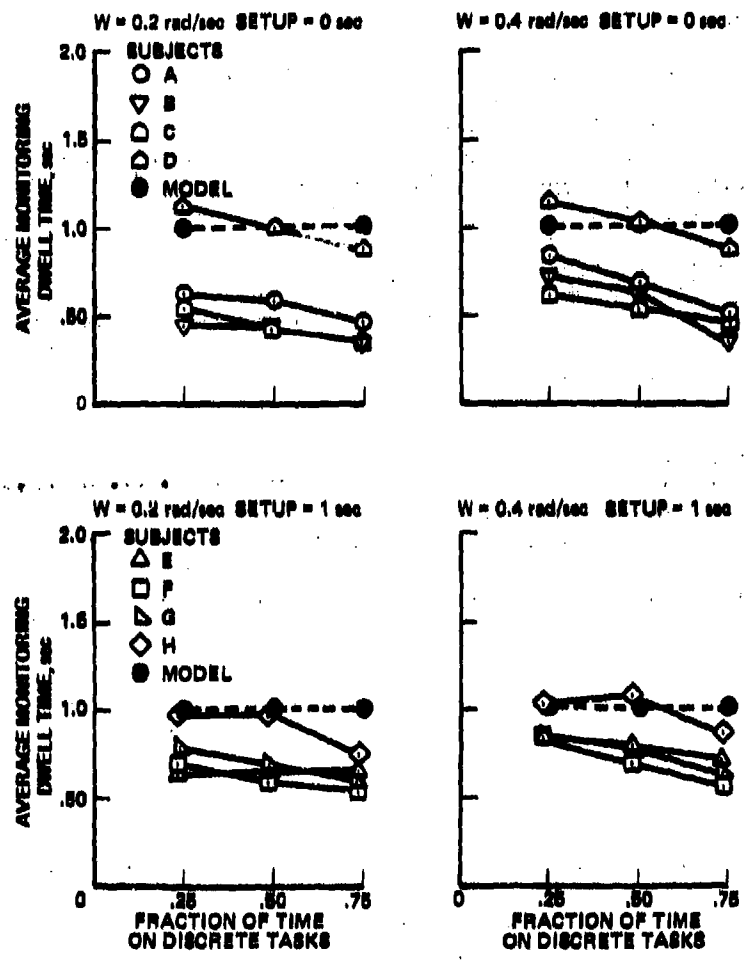


Figure 7. The average dwell time on the monitoring display for dwell times less than 1.5 sec long. The model's minimum dwell time was fixed — usually at 1 sec.

An ANOVA on the fraction-of-hits data showed that all main effects were significantly different, at a minimum of the 0.01 level. All two-way interactions were significant at a minimum of the 0.05 level. Only the three-way interaction was insignificant. An identical ANOVA was run on the difference between the subjects' and the model's fraction of hits data. None of the main effects or interactions were significant at the 0.05 level.

Figures 8, 9, and 10 show summary statistics on the strategy of the subjects and the model. The average number of switches to the discrete task was higher for both subjects and model when there was no setup cost or task-switching penalty. The variability among the subjects was greater when there was no setup cost. Setup cost reduces the benefit of reducing the minimum dwell time on the monitoring task. The average number of switches was only slightly affected by the fraction of time that could be devoted to monitoring. It is interesting, however, that both the model and most of the subjects made the greatest number of switches when they could spend only half of their time monitoring. Figures 9 and 10 show that both the mean and the standard deviation of the discrete task dwell time increased sharply when less time was spent monitoring. The average discrete task dwell time also increased when there was a setup cost.

Figure 11 shows the average fraction of time that had to be spent on the discrete task when the time to the next waypoint was 30 sec. If the data fell on the diagonal line, the subject had done half of the discrete task in half the time allowed to do it. If the data fell above the diagonal, the subject was postponing the discrete tasks until later in the waypoint segment. With no setup cost, subjects A, B, and C tended to postpone the discrete tasks when only 25% of the time was spent monitoring. With the 1-sec setup cost, there was a larger tendency to postpone the discrete tasks.

The above data and other summary measures on performance and strategy show good agreement between the subjects and the model. Figures 12 and 13 show the results of comparing each subject's decision to switch to the discrete task with the decision the model would have made in the same situation. The only subject-decisions excluded from this comparison were those in which the amount of time remaining to be spent on the discrete task was within 1 sec of the time remaining until the next waypoint was reached. Figure 12 shows the fraction of time that the model agreed with the subject's decision to switch to the discrete task. The agreement tends to decrease as the task becomes more difficult, either because the process bandwidth increases or because the setup cost increases. Note that with no setup cost, subject D had a good agreement with the model decisions to switch away but had the worst performance. Subject D typically switched to the discrete task at good times but his discrete task dwell times were too long. With the setup cost, subject G performed worst and had the lowest agreement with the model.

Next, the correlation coefficients were computed between discrete task dwell times for the subject and the model. These data are shown in figure 13 for each subject and for the 12 experimental conditions. The correlation coefficients range from 0.13 to 0.79, with most near 0.5. The correlations were generally higher with the lower bandwidth process. On that more microscopic level, the match between the model and the data was not as good.

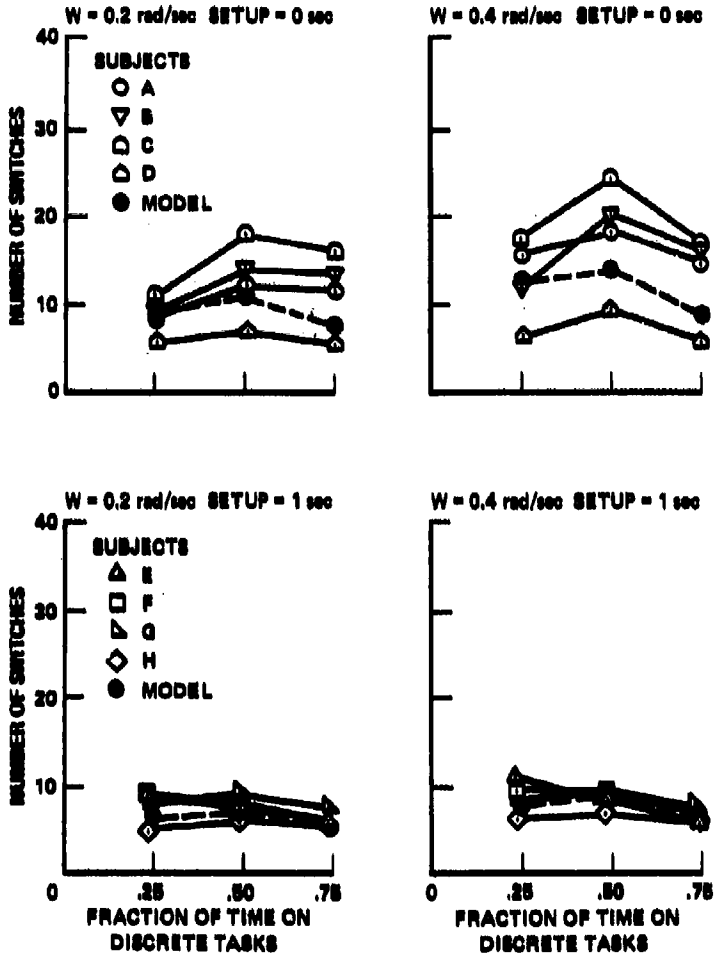


Figure 8. The average number of switches from the monitoring display to the discrete task per waypoint segment.

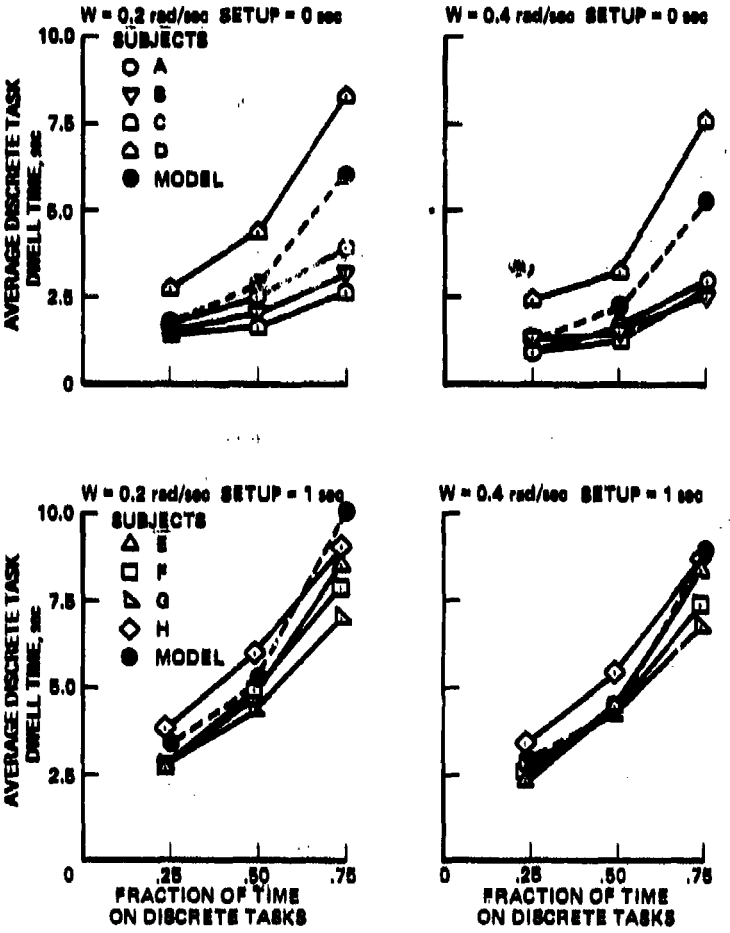


Figure 9. The average dwell time on the discrete task.

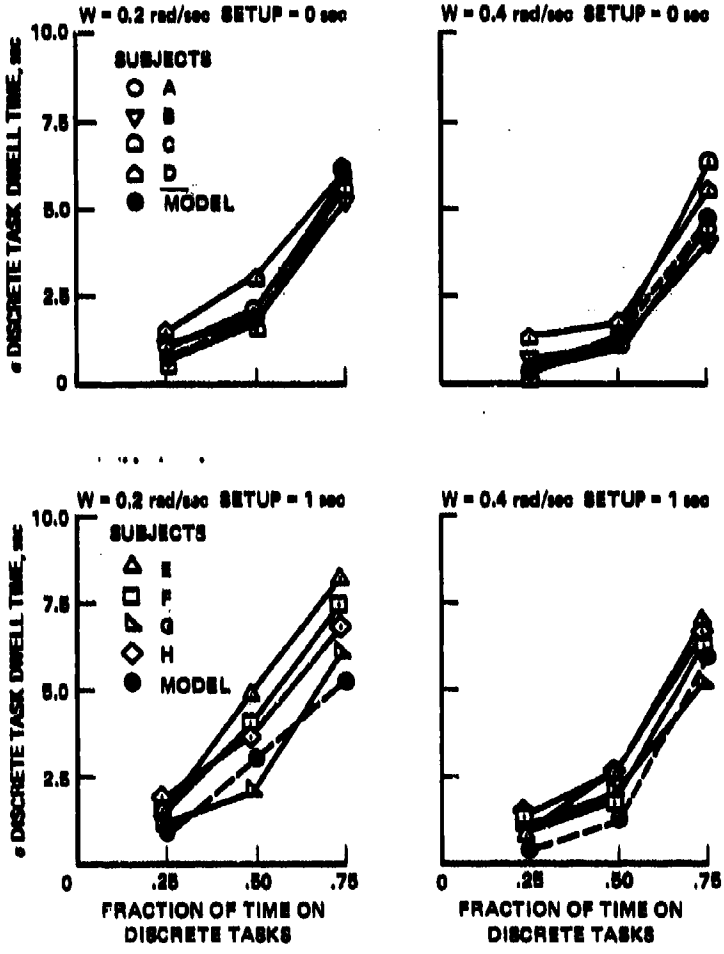


Figure 10. The standard deviation of the dwell times on the discrete task.

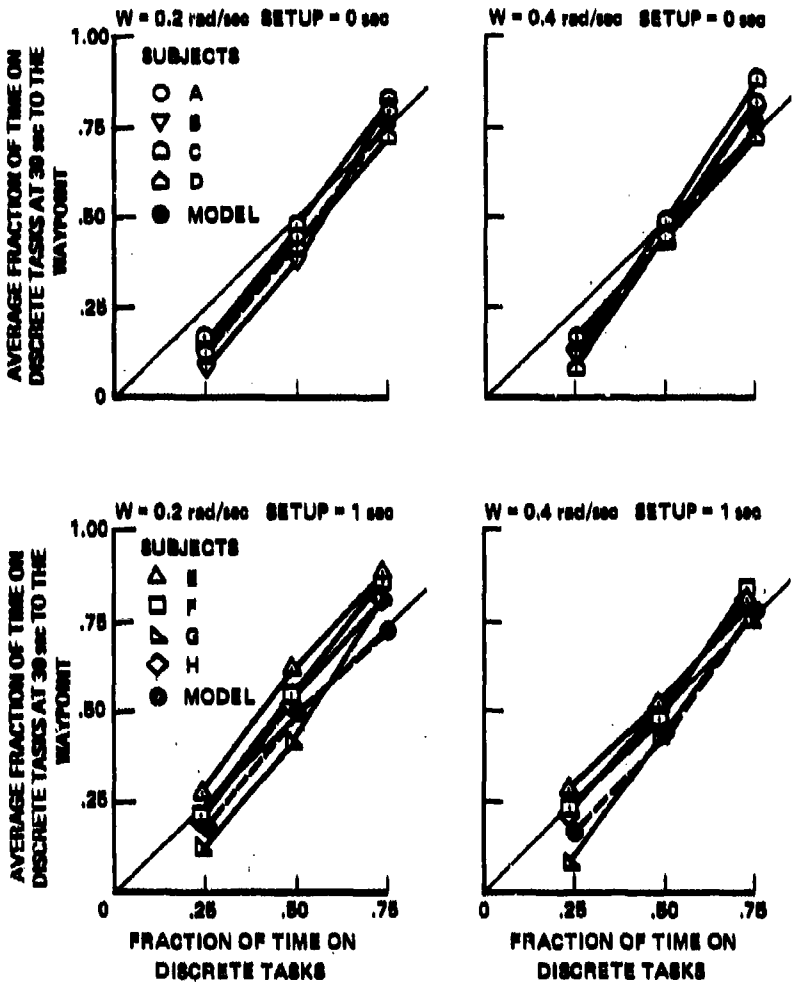


Figure 11. The fraction of the remaining time which must be spent on the discrete task when the next waypoint is 30 sec away.

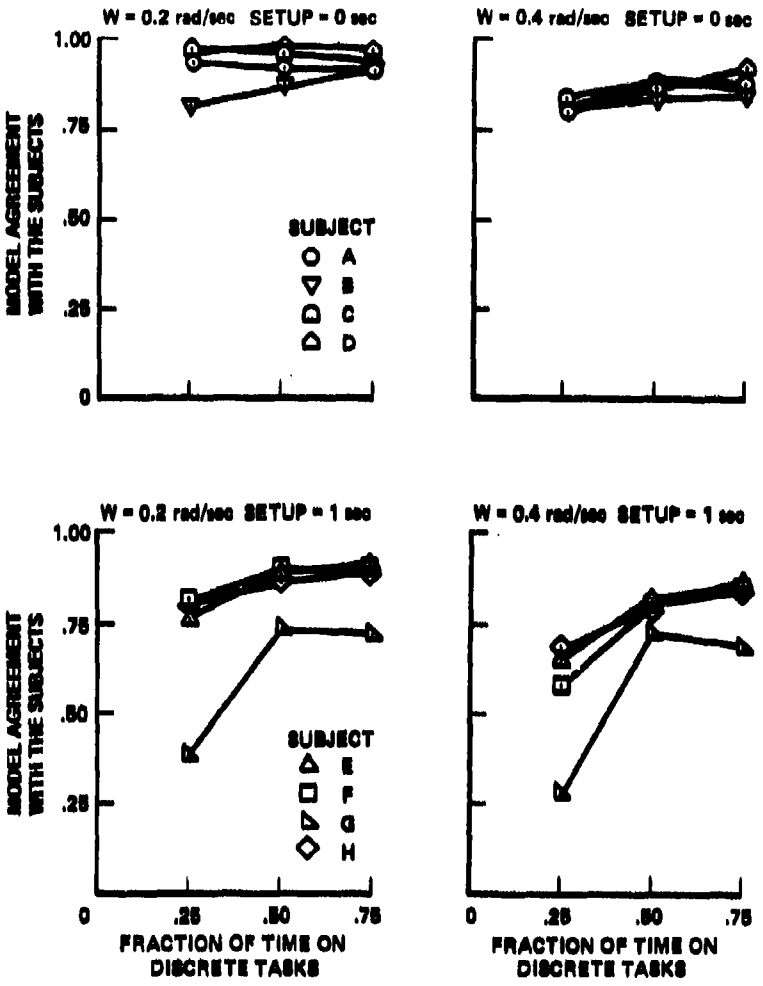


Figure 12. The fraction of decisions in which the model agreed with the subjects decision to switch to the discrete task.

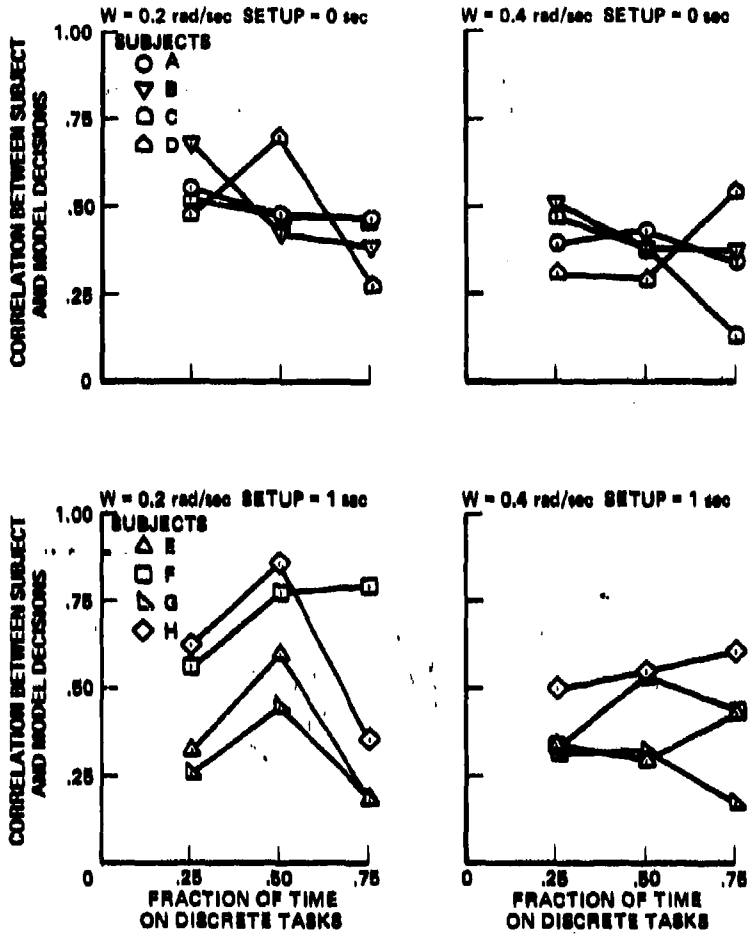


Figure 13. The correlation between each subject's discrete task dwell times and the model's decision in the identical situation.

In order to describe the subjects' time-sharing strategies, multiple linear regression analyses were performed to determine functional relationships between the subjects' decisions and key task variables at the time the decisions were made. The dependent variable was the discrete task dwell time in seconds (D). The independent variables were the absolute magnitude of the display error (e), the error rate ($\dot{e}$), and the fraction of the remaining time that had to be spent on the discrete task (F) at the time each decision was made, and the squares and cross-products of those three terms. In order to eliminate extreme conditions, decisions were eliminated from the analyses if any of the following conditions were met: (1) $D < 0.3$ sec; (2) $e > 0.93$ dots; and (3) $F < 0.1$ or $F > 0.9$. Data for each subject and the two process bandwidths were analyzed separately but data for the three fraction-of-time-on-discrete-task conditions and the two replications were combined. In order to allow comparisons between the subjects and the model, identical analyses were done on the model's decisions.

The regression coefficients that were significantly different from zero, the multiple correlation coefficients, and the correlation between the decisions of the model and subjects are listed in table 1. Many more of the regression coefficients were significantly different from zero for the model than for the subjects. The magnitude of the coefficients for the model were also almost always larger. One reason for the higher number of significant coefficients for the model is, of course, that the model is completely consistent. It always made the same decision for a given value of e, $\dot{e}$, or F. The lower magnitude of the coefficient for the subjects suggests that they made less use of the available information. In general, the subjects' strategies appeared to be to switch away from the monitoring display for a relatively constant length of time. Their decisions were a much weaker function of the display state and the fraction of time that had to be spent on the discrete task.

On the condition with no setup cost, subjects A, B, and C appear to be quite similar to each other but distinctly different from subject D. Subject D's intercepts were 5.2 and 6.0 sec, which was about 3 times higher than the intercepts for subjects A, B, and C. Subject D's decisions were correlated most highly with F and F^2 and only slightly with the display state. Dependence on F is necessary to finish the discrete task on time but the display state provides the information that is necessary to get a high monitoring score. It is somewhat ironic that although subject D's decisions correlated most highly with the model he got the lowest monitoring score.

The decisions of subjects A, B, and C were more strongly correlated with the display state and those subjects got higher monitoring scores. The significant regression coefficients were usually with e, $\dot{e}$, and $e\dot{e}$. A negative e coefficient implies a strategy that switched away for shorter times when the error was close to the out-of-tolerance limit. A negative $\dot{e}$ coefficient implies that decisions were shorter when the error signal was moving rapidly in any direction. A negative $e\dot{e}$ coefficient implies that error-rate information affected decisions most strongly when the error signal was close to the tolerance limit.

On the setup cost conditions, the intercepts were generally larger. The subjects' decisions were also more correlated with F^2. Again, the subjects

TABLE 1

REGRESSION COEFFICIENTS AND RELATED STATISTICS FOR THE BEST LINEAR RELATIONSHIP BETWEEN THE SUBJECT'S DISCRETE TASK DWELL TIME (D) AND THE FOLLOWING INDEPENDENT VARIABLES: (1) THE ABSOLUTE MAGNITUDE OF THE DISPLAY ERROR (e); (2) CORRESPONDING VALUE OF THE ERROR RATE ($\dot{e}$); (3) THE FRACTION OF THE REMAINING TIME THAT MUST BE SPENT ON THE DISCRETE TASK (F); AND (4) THE SQUARES AND CROSS-PRODUCTS OF THESE TERMS.

BANDWIDTH = 0.2 rad/sec, SETUP = 0 sec

Variable	Model	Subject A	B	C	D
Intercept	1.4	2.1	2.1	1.4	5.2
e	2.6	---	---	0.5	---
$\dot{e}$	4.6	---	---	---	---
F	---	2.8	1.8	0.8	-9.1
e^2	-2.3	-1.8	-1.4	-1.0	---
$\dot{e}^2$	-5.0	-3.9	-3.8	-0.9	---
F^2	12.7	---	---	---	21.5
$e\dot{e}$	-5.7	-1.7	-3.6	-1.2	-6.4
eF	-6.2	---	---	-0.6	---
$\dot{e}$F	-15.9	---	---	---	---
N	215	496	584	772	275
Stan Err	.85	.82	1.08	.55	2.22
Mult R	.95	.71	.55	.57	.75
R with Model	1.00	.43	.35	.39	.66

BANDWIDTH = 0.4 rad/sec, SETUP = 0 sec

Variable	Model	Subject A	B	C	D
Intercept	1.2	0.6	1.8	0.8	6.0
e	2.7	0.7	---	---	---
$\dot{e}$	2.3	---	---	---	---
F	---	3.1	---	1.7	-20.6
e^2	-3.6	-1.0	-1.2	-0.7	-3.4
$\dot{e}^2$	-3.2	-0.5	-0.7	-0.3	---
F^2	9.7	---	1.4	-0.8	31.6
$e\dot{e}$	-6.1	-0.5	-0.9	-0.5	---
eF	-3.4	-0.8	---	---	---
$\dot{e}$F	-4.7	---	---	---	---
N	278	741	742	909	306
Stan Err	.56	.58	.80	.36	2.29
Mult R	.95	.70	.44	.60	.76
R with Model	1.00	.55	.31	.41	.64

BANDWIDTH = 0.2 rad/sec, SETUP = 1 sec

Variable	Model	Subject E	F	G	H
Intercept	4.5	2.4	2.7	4.5	6.8
e	2.7	---	---	-3.3	-7.0
$\dot{e}$	7.2	---	---	---	---
F	-7.4	---	---	---	---
e^2	-6.3	-2.4	-3.0	---	---
$\dot{e}^2$	-31.6	-4.5	-11.6	---	-21.9
F^2	28.6	10.0	---	9.7	8.0
$e\dot{e}$	-27.8	-8.4	-12.5	---	-10.5
eF	-4.7	---	---	5.3	5.3
$\dot{e}$F	-12.3	---	---	---	---
N	147	328	312	257	229
Stan Err	1.05	1.82	1.67	2.36	2.17
Mult R	.97	.67	.70	.64	.72
R with Model	1.00	.64	.71	.56	.73

BANDWIDTH = 0.4 rad/sec, SETUP = 1 sec

Variable	Model	Subject E	F	G	H
Intercept	2.7	0.9	2.4	3.8	4.3
e	5.7	---	---	---	---
$\dot{e}$	6.8	---	---	---	3.2
F	-6.6	---	---	---	---
e^2	-3.8	-4.2	---	---	---
$\dot{e}^2$	-5.6	---	-1.9	---	-4.6
F^2	25.7	7.5	7.3	10.6	15.0
$e\dot{e}$	-8.3	---	-4.0	---	-6.9
eF	-10.4	---	---	---	-5.2
$\dot{e}$F	-15.6	---	---	---	---
N	186	338	333	249	235
Stan Err	.71	2.17	1.43	2.83	2.32
Mult R	.97	.67	.76	.54	.72
R with Model	1.00	.56	.74	.47	.71

(E, F, and H) who performed the best under those conditions were the ones whose decisions were the most strongly correlated with the display state.

After each run, the subjects rated the task on the bipolar adjective pairs shown in table 2. The fraction of time that had to be devoted to discrete tasks had the strongest influence on the subjective ratings. As that parameter increased, the subjects generally agreed that the task became more "demanding," "hard to learn," "annoying," "complex," "unpredictable and difficult." The subjects rated the task as being "interesting" and requiring "skill" with the exception of subject G on the higher bandwidth task. The tasks were generally ranked as being more "frustrating" by the subjects who had the experimental conditions with the 1-sec setup cost.

TABLE 2
BIPOLAR ADJECTIVE PAIRS USED TO RATE EACH EXPERIMENTAL CONDITION

Demanding	Undemanding
Hard to learn	Easy to learn
Confusing	Clear
Surprising	Routine
Annoying	Pleasant
Passive	Active
Complex	Simple
Incuitive	Rational
Uninteresting	Interesting
Subtle	Obvious
Frustrating	Soothing
Unpredictable	Predictable
No skill required	Skill required
Easy	Difficult

CONCLUDING REMARKS

In this paper the general problem of time-sharing attention between instrument monitoring and other duties was described and a simulation model based on the internal model concept was presented for that kind of task. Monitoring-performance of the model was presented in terms of the fraction of out-of-tolerance errors observed as a function of the fraction of time spent on discrete tasks. Performance was shown to be sensitive to process bandwidth and discrete task setup cost. Performance and strategy were not sensitive to errors in the simulated operator's internal model or to the lack of rate information in the display. Performance was sensitive to the minimum dwell time that the model spent on the monitoring display.

An experiment was conducted in which eight subjects monitored the output of a second-order process. All of the independent variables — fraction of time available to monitor, process bandwidth, and discrete task setup cost — significantly affected monitoring performance. Summary measures of the models performance and strategy agreed well with the subjects' data. The fit improved further when the model was rerun with a dwell time closer to that used by the subjects. A decision-by-decision comparison between the model and

each of the subjects indicated that decisions of the subjects were not as strongly correlated with the display state and the discrete task as the model. The fact that the subjects did not act as decisively as the model on the relevant information might be attributable to risk-averse behavior.

Future work will apply these results to the analyses and interpretation of an experiment in which subjects will be required to divide their attention between a monitoring activity and a more realistic discrete task in which they will have to key in numeric data.

REFERENCES

1. Palmer, E. A., Models for Interrupted Monitoring of a Stochastic Process, NASA TM-78,453, Oct. 1977.

2. Smallwood, R. D., Internal Models and the Human Instrument Monitor, IEEE Trans. on Human Factors in Electronics, Vol. HFE-8, No. 3, Sept. 1967, pp. 181-187.

3. Palmer, E. A., Interrupted Monitoring of a Stochastic Process, Proceedings of the 13th Annual Conference on Manual Control, MIT, 1977.

USING MODELING TO DETERMINE MEASURES OF NON-PRODUCTIVE CONTROL ACTIVITY

D. W. REPPERGER A. M. JUNKER
AEROSPACE MEDICAL RESEARCH LABORATORY
WRIGHT-PATTERSON AIR FORCE BASE, OHIO 45433

ABSTRACT

This study will consider a measure of human control effort in a tracking task which is not productive in the minimization of the closed loop tracking error. To determine this measure of control activity, a Kalman filter is used in this modeling approach to estimate the stick output of the human. Using the Kalman filter estimate, the stick signal is decomposed into two components. One component of the stick signal is correlated with the closed loop error signal; the second component is orthogonal to the closed loop error signal. Using phasor representation, these stick components are multiplied by the system to be controlled (as a phasor) and reversed in sign to see the effect of the stick output on the error signal. After rotating and translating the component vectors, the true stick power orthogonal to the error signal is determined.

In studies related to learning, it is hypothesized that if large amounts of control activity are not productive in the tracking task, the subject may be in the preliminary stage of a learning process. As the subject becomes more acquainted with the task, his measure of control activity correlated with the reduction of closed loop error should increase. This aspect of learning provides an interesting interpretation of human response behavior. To relate this modeling procedure to empirical data, a motion cue time delay experiment was considered. In the motion study, all subjects tracked the same vehicle dynamics but with various amounts of motion cue information as a result of delaying the motion response of the simulator by different amounts. For increased delays, tracking performance was worse and the subjects rated the task more difficult. This empirical data base provides a unique opportunity to investigate several possible measures of human control effort not productive in the tracking task. This measure of control activity is studied with respect to the learning conditions in this experiment.

Introduction

Investigations of human operator modeling during learning has been of interest in many studies. In [2, 3] an investigation of how performance changes due to the learning in the tracking task was conducted.

In [1] a Kalman filter was used to determine measures of human invariance in the tracking task. It can be shown that these results can be extended to include measures of orthogonality in tracking. It is the purpose of this study to extend the previous results and to investigate how these measures change as a function of learning. This provides a unique opportunity to study human behavior as a function of how the human learns the tracking task.

A Modeling Approach to Study Learning

Figure (1) illustrates the type of modeling approach that will be used here. With respect to Figure (1), a post-experimental analysis of the input-output time series associated with the human operator will be conducted. The input time series is assumed to be functions of e(t), the closed loop system error. The output of the human u(t) will be modeled via the Kalman filter program. Figure (2) illustrates the updating procedure to obtain the Kalman filter parameters.

With reference to Figure (2), the inputs to the model are functions of the time series e(t) (the displayed error signal). The purpose of this modeling approach is to choose model parameters such that the model's output $\hat{x}(t)$ is an accurate representation of the measured stick response of the human. The measure of modeling accuracy is expressed in the residuals or output modeling error $\nu(t)$ which satisfies:

$$\nu(t) = u(t) - \hat{x}(t) \tag{1}$$

If the model is appropriately fitted to the data, then $\nu(t)$ should be a random white process which satisfies:

$$\text{mean } [\ \nu(t)\ ] = E[\nu(t)] = 0 \tag{2}$$

$$\text{var } [\nu(t)] = E[\nu(t)\nu^T(\tau)] = R\ \delta(t-\tau) \tag{3}$$

It will be necessary in the subsequent analysis to test $\nu(t)$ for whiteness and determine R of equation (3). If $\nu(t)$ is a random white process, then the expected value of the model is equal to the expected value of the human's output.

In Figure (2), it is desired to update the model parameters in such a way that the innovations sequence $\nu(t)$ is a white, random process. The method of updating the parameters is based on an algorithm [4] which is actually a maximum liklihood procedure. In this manner a unique value of the optimal gain can be determined which maximizes the probability density function of the structure of the assumed model based on all the available data points.

With reference to Figure (2), a Quasi-linear human operator model will be assumed in this approach. The human's input-output characteristics are described by:

$$\dot{x}(t) = A\,x(t) + B\,\bar{e}(t) + \zeta(t) \tag{4}$$

where x(t) is a 2x1 dimensional column vector with two components defined as follows:

$$x_1(t) \triangleq u(t) \tag{5a}$$

$$x_2(t) \triangleq \frac{d}{dt}\,u(t) \tag{5b}$$

$\zeta(t)$ is a 2x1 dimensional white Gaussian noise source with mean zero and covariance Q. The 2x2 dimensional matrices A and B are to be determined. The 2x1 column vector $\bar{e}(t)$ satisfies:

$$\bar{e}(t) = \begin{bmatrix} e(t-\tau) \\ \dot{e}(t-\tau) \end{bmatrix} \tag{6}$$

The empirical time series $\bar{e}(t)$ is determined from values of the empirical data e(t) and by using a delay of τ units. $\dot{e}(t-\tau)$ is obtained via a digital filter processed on the data and τ is approximated from the Bode plot of the transfer characteristics of the human operator(c.f. Figure (3)). Using mean values of equation (4), a transfer function representation is desired of the form:

$$Y_p(s) = \frac{\bar{U}(s)}{E(s)} = \frac{d(s+a)}{(s+b)\,(s+c)}\,e^{-\tau s} \tag{7}$$

where $\bar{U} = E[\,u(t)]$,

d is the d.c. gain, a zero occurs at a radians/second, and the poles are at b and c radians/second represent typical second order characteristics of the human. Second order characteristics appear to be sufficient to describe a human operator for the modeling purposes [5, 6]. Taking the Laplace transform of the mean value of equation (4) shows the parameters d, a, b, and c must satisfy the following relationship:

$$Y_p(s) = \frac{\bar{U}(s)}{E(s)} = \frac{d(s+a)\,e^{-\tau s}}{(s+b)\,(s+c)} = [1,0]\,[sI-A]^{-1}B\begin{bmatrix} 1 \\ s \end{bmatrix} e^{-\tau s} \tag{8}$$

Appendix (A) illustrates the manner in which the A and B matrices are constructed such that the representation in equation (7) can be chosen from a Bode plot and the relationship (8) will be satisfied. Figure (3) illustrates a typical Bode plot averaged across replications in the frequency domain of the describing function of the human for one experimental condition of the data considered in this paper. In this paper the assumption will be made that the spectra of $\bar{u}(j\omega)$ can be approximated by the spectra of $u(j\omega)$. From inspection of equation (4) it is seen that the two time series differ by a low pass noise process. Since the describing function fit is determined in the low frequency region, the dynamics of equation (7) are chosen below the poles of $[SI-A]^{-1}$ which is consistent with this assumption [7].

With reference to equation (4), it is desired to estimate x(t). To achieve this goal, a Kalman filter will be used.The measurement equation to complement equation (4) is given by:

$$y(t) = u(t) + n(t) = Hx(t) + n(t) \tag{9}$$

which implicitly assumes that the human has available linear combinations of the error signal and the rate of the error signal through his various sensor processes (this is easily seen to be true by solving equation (4) and substituting x(t) into (9), c.f. Appendix A).

In order to apply this identification procedure, the linear differential equation (4) will be used with the measurement equation (9). The time series e(t) is delayed τ units and the vector $\bar{e}(t)$ is determined and used in the identification procedure. Under these assumptions, the Kalman filter equation is then specified by:

$$\frac{d}{dt}\hat{x}(t) = A\,\hat{x}(t) + K\,v(t) + B\,\bar{e}(t) \tag{10}$$

where $\hat{x}(t)$ is a 2x1 column vector which represents the best linear least squares (minimum variance) estimate of $\hat{x}(t)$ based on the available measurements $y(t)$. $\hat{x}(t_o)$ satisfies an unbiased condition:

$$\hat{x}(t_o) = 0 \tag{11}$$

The Kalman gain K satisfies:

$$K = P\,H^T\,R^{-1} \tag{12}$$

where P is the steady state solution to the following Riccati equation:

$$\dot{P} = PA + A^TP - PH^TR^{-1}HP + Q \tag{13}$$

The optimal gain is the principal part of the discrete Kalman filter model which is described by:

$$\hat{x}_{i+1/i} = \phi\, \hat{x}_{i/i} + \int_0^{\Delta t} e^{A\tau} d\tau B \mathrm{col}[e(t), \dot{e}(t)] \tag{14}$$

$$\hat{x}_{i/i} = \hat{x}_{i/i-1} + K_o\, [y_i - H\, \hat{x}_{i/i-1}] \tag{15}$$

Where $\hat{x}_{i/i}$ is the minimum variance estimate of the human's stick response. The matrix ϕ is the discrete transition matrix associated with the human's transfer function determined as follows:

$$\phi = e^{A\Delta t}$$

where $\Delta t = .04$ seconds (the sampling rate) and the matrix A is determined from the relationship (8) (see Appendix A).

The matrix H in equation (9) is specified by $H=[1,0]$. The Kalman gain K_o satisfies:

$$K_o = P\, H^T (HPH^T + R)^{-1} \tag{16}$$

$$P = \phi\, [P - PH^T (HPH^T + R)^{-1}\, HP] \phi^T + Q \tag{17}$$

Where the covariance matrices Q and R describe the human's uncertainty in the tracking task. The manner of obtaining the Q and R matrices is based on the algorithm in [4]. Initial matrix values denoted as Q_o and P_o are chosen. In order to establish the updating rule, it is necessary to define the sample covariance function:

$$\text{Let:} \quad \hat{C}_k = \frac{1}{N} \sum_{i=k}^{N} \nu_i \nu_{i-k}^T \tag{18}$$

is a sample covariance function. The matrices R and Q are now updated [4] via:

$$R_k = \hat{C}_o - H(P_k H^T) \tag{19}$$

$$\text{where} \quad P_k H^T = K_o\, \hat{C}_o + A^* \begin{bmatrix} \hat{C}_1 \\ \hat{C}_2 \end{bmatrix} \tag{20}$$

$$\text{where} \quad A^* = (A^{-T} A)^{-1} A^{-T} \tag{21}$$

and $\bar{A} = \begin{bmatrix} H\phi \\ H\phi\ (I-K_oH)\phi \end{bmatrix}$ (22)

and finally Q is determined via:

$$Q_k = P - \phi KR_kK^T\phi^T - \phi(I-KH)P(I-KH)^T\phi^T \quad (23)$$

This algorithm has been shown to converge [4] and is equivalent to maximizing the log-likelihood function of the model structure conditioned on the date. For this problem the log-likelihood functional is specified by:

$$J(\Theta) = -(1/2) \sum_{i=1}^{N} [v^T(t_i,\Theta)\ \hat{C}_o^{-1} v(t_i,\Theta) + \log |\hat{C}_o| \ ] \quad (24)$$

Where N is the number of samples. The t_i in equation (24) refers to the time instants where the N data samples occur. The choice of parameters $\Theta=\Theta$ (Q,R) which maximizes J is equivalent to maximizing the density function of the parametric form conditioned on the data.

The final validation of this modeling effort is the need to test the residuals for whiteness. To accomplish this goal the normalized auto correlation function $\hat{\rho}_k$ is computed as follows:

$$\hat{\rho}_k = \frac{\hat{C}_k}{\hat{C}_o} \quad (25)$$

The test of whiteness of the residuals is a 95% whiteness test on $\hat{\rho}_k$. The 95% confidence limits for ρ_k are $1.96/\sqrt{N}$ where N is the number of samples. The band $\pm 1.96/\sqrt{N}$ is constructed about zero. If less than 5% of the sample points lie outside the band, the sequence is white. If more than 5% of the sample points lie outside the band, then a significant correlation exists in the residuals and the sequence is not white. Figure (4) illustrates the sample auto-correlation function obtained here from the data after the residuals have been whitened via this algorithm. A description of the empirical data base is presented next.

The Time Delay Experiment

In an effort to study the effects on learning and performance of delaying motion cues to the human operator, a time delay experiment was conducted with respect to studying motion effects on the human operator. Figure (5) illustrates the experiment considered here. A disturbance type input f(t) was chosen for this experiment since this is where the motion effects are most pronounced. Five experimental condition were considered. The static condition consisted of tracking the disturbance input for a plant specified by:

$$H(s) = \frac{14}{(s)} \ \frac{5}{(s+5)} \cdot \frac{19}{(s+19)} \ e^{-.065s} \quad (26)$$

i.e. due to the analog-digital interface, there existed a delay of 65 msec in processing the feedback signal to the display in the static condition. With reference to Figure (5), four motion conditions were considered. The synchronous motion condition was accomplished by rotating the tracker in the roll axis due to his stick movements. The rotation was delayed 65 msec until the display and physical rotation coincided. Three other motion conditions were also considered. The motion loop was delayed 80 msec, 200 msec, and 300 msec to produce distorted motion cues to the human operator. These delays were in addition to the 65 msec due to the analog-digital processing time.

Table I was constructed which is based on ensemble mean scores of subjects involved in the experiment.

Table I Asymptotic Training Scores

EXP CONDITION	ms ASYMPTOTIC $_2$ ERROR (deg^2)
STATIC	32.9
SYNCHRONOUS MOTION	7.5
MOTION 80ms DELAY	11.3
MOTION 200ms DELAY	27.3
MOTION 300ms DELAY	39.6

From Table I, one can study the effect of motion on performance by comparing the synchronous motion score to the static score. Also a delay of 200 msec to 300 msec produce a significantly degrading effect such that motion in this mode is no better than the static score. As a result the improvement due to motion is cancelled by a delay of approximately 250 msec.

The interesting aspect of this experiment can be seen in the learning period scores. Figure (6) illustrates these values for the subjects Tom and Cathy.

It is from this type of learning curve that it is desired to study this effect in conjunction with the modeling procedure. First the subject Tom on days 2, 5, and 9 is considered for the static condition. The subject Cathy is chosen for the 200 msec delay condition during days numbered 2, 5, and 10. It is from this type of data base that we wish to determine measures of human response orthogonal to the task or goal. We next discuss how to determine these orthogonal measures.

Determination of Orthogonal Measures

With reference to Figure (7), the man-machine problem is viewed in a phasor framework. Denoting $E(j\omega)$ as the phasor representation for the error vector, Figure (8) illustrates the projection of the plant's output

$X(j\omega)$ into two components $X_1(j\omega)$ and $X_2(j\omega)$. For simplicity we denote as $x(t)$ the plant's output due only to the human control action $u(t)$ (see bottom of p. 9).

<u>if</u> $$x_1(t) + x_2(t) = x(t) \tag{27}$$

<u>and</u> $$E\{x_1(t)\, e(\tau)\} = 0 \tag{28}$$

Where $x_1(t)$ represents that part of the plant's output orthogonal to $e(t)$. Since $x_1(t)$ represents that portion of the plant's output not useful in tracking, the question is raised as to what portion of the human response vector $u(t)$ gave rise to $x_1(t)$ and $x_2(t)$. In other words:

$$u(t) = u_1(t) + u_2(t) \tag{29}$$

Where $u_1(t)$ is that portion of the human's stick response that produces $x_1(t)$. Figure (9) illustrates the total phasor diagram. In Figure (9), the stick response phasor $U(j\omega)$ is decomposed into components u_b and u_a which are orthogonal and correlated, respectively, with the error phasor $E(j\omega)$. Since:

$$X(j\omega) = H\, U \angle(\omega_1 + \theta_a) \tag{30}$$

which represents the plant output phasor $X(j\omega)$. It is of interest here to determine what component of the human stick response gives rise to $X_1(j\omega)$ where:

$$X(j\omega) = X_1(j\omega) + X_2(j\omega) \tag{31}$$

and $$E\{e(t)\, x_1(\tau)\} = 0 \tag{32}$$

in other words, if: $$u(t) = u_1(t) + u_2(t) \tag{33}$$

where $u_1(t)$ is that part of the stick output which produces $x_1(t)$ which is orthogonal to the error. This is a measure of non-productivity control activity not directly applicable to improving the task at hand. With reference to Appendix A, $u_1(t)$ can be determined as follows:

$$u_1(t) = H \int_{t_o}^{t} \phi(t,s) K(s) v(s)\, ds + Hv(t) + \eta(t) \tag{34}$$

and it can be shown:

$$u_2(t) = H \int_{t_o}^{t} \phi(t,s) B\, \bar{e}(s)\, ds \tag{35}$$

To determine the stick power not productive in the tracking task, a spectra analysis procedure is combined with the Kalman filter program as follows:

(1) Using the Kalman filter program, the stick output $u(t)$ is decomposed into one component correlated with the error vector ($u_a(t)$) and a second component ($u_b(t)$) orthogonal to the error vector.

(2) A spectra analysis is conducted of $u_a(j\omega)$ and $u_b(j\omega)$ to determine the power at each frequency.

(3) Using the transfer function of the plant, the magnitude of the plant transfer function and phase shift at each frequency is determined.

(4) Using the spectra of $u_a(j\omega)$ and $u_b(j\omega)$ the phasors are shifted in Figure (9) and $X_1(j\omega)$ and $X_2(j\omega)$ are determined.*

From this procedure, the stick output $u_1(t)$ which produces no productive effect on the error signal can be determined for each experimental condition. It is emphasized that the Kalman filter is required to initially generate u_a and u_b which is used to determine x_1 and x_2. The next section discusses the results obtained from this empirical data base.

Results of This Analysis

Table II and III illustrate the results of this analysis for the subjects Tom and Cathy. The term S/N is calculated as follows:

Table II

STATIC - TOM

DAY	EXP	ERROR	ρ_{11}	v^2	S/N	R_z
2	76	47.00	73	8.8185	87.49	.0044
5	240	35.42	35	8.0841	95.543	.0037
9	473	33.65	54	7.8281	115.18	.0034

Table III

200.0 ms DELAY-CATHY

DAY	EXP	ERROR	ρ_{11}	v^2	S/N	R_z
2	142	59.11	74	17.5102	104.16	.0072
5	245	35.63	95	11.94	105.22	.0062
10	513	30.25	87	10.548	117.494	.0055

* It is noted that in this procedure we ignor the plant's response to the disturbance input $f_{Disturbance}$. The true plant output has a component due to u(t) and also to $f_{Disturbance}$. We wish in this study to only analyze the effect on the error signal e(t) of the response characteristics that are due to u(t).

$$S/N \approx \frac{\sum_{i=1}^{N} (\hat{u}(t_i))^2}{\sum_{i=1}^{N} v^2(t_i)} \qquad (36)$$

i.e. it is a measure of signal to noise characteristics of the human operator. The term $\hat{u}$ is a measure of the correlation of the human's output with the input vector e(t). The residuals v(t) are the identified residuals from this modeling process. The quantity R_z is the sample output noise covariance, i.e.

$$R_z = \frac{1}{N} \sum_{i=1}^{N} v^2(t_i) \qquad (37)$$

Which is a measure of human randomness. The term ρ_{11} is the number of points in the sample (N=1200) outside the range of whiteness by the finite data whiteness test illustrated in Figure (4). For a 95% whiteness test, the F ratio test requires that no more than (.05) (1200)=60 samples of the normalized auto correlation function be greater than 1.96/ $\sqrt{1200}$. This is displayed in Figure (4).

The results of Tables II and III illustrate some interesting results relating learning and tracking performance. In the static case for the subject Tom, as learning improved (days 2, 5, and 9), the values of the residuals (Σv^2) decreased indicating more certainty in tracking. Correspondly, the S/N ratio improved indicating a more certain tracking strategy. This result concurs with increasing pilot gain [2, 3] as the subjects became more acquainted with the tracking task.

For the 200 msec delay condition, the subject Cathy exhibited the same trends in tracking performance as the asymptotic levels were reached. The sum of the squared residuals continuously decreased and an improvement in S/N resulted. One could use the measure of S/N developed here for the relative tracking accuracy of a subject. It is noted that this measure is obtained in the time domain as contrasted to frequency domain measures used in other approaches.

As a measure of learning or efficiency in tracking, the following measure was proposed:

$$J = \frac{\sum_{i=1}^{\bar{n}} \text{Power } u_1(\omega_1)}{\sum_{i=1}^{\bar{n}} \text{Power in } u(\omega_1)} \qquad (38)$$

where $\bar{n}$ denotes the number of frequencies in which enough power exists of sufficient magnitude to be included in the estimate. From this measure, if J=1, then the entire stick response is uncorrelated with reduction of the error. If J=0, then the total stick output is productive in reducing the error. Table V illustrates the values of J obtained here:

Table V

SUBJECT	DAY	ERROR SCORE	J CALCULATED
TOM	2	47.00	.40
TOM	5	35.42	.32
TOM	9	33.65	.27
CATHY	2	59.11	.45
CATHY	5	35.63	.34
CATHY	10	30.25	.29

It is noted that for the static case, as the subject Tom becomes more familiar with the tracking task, J decreases to a more efficient value. Likewise in the 200 msec delay condition, the subject Cathy, also decreased her J value as learning improved. This merely reflects the percentage of stick power useful for this tracking task. The remaining stick power (.27 and .29 for Tom and Cathy, respectively) is necessary to obtain monitoring and motion or vision cues from the closed loop system. This certain amount of control activity not productive in tracking is beneficial to the human operator for monitoring and decision making.

Summary and Conclusions

A study is being conducted to measure how much of human's output is not productive in tracking. A ratio is obtained on the percent of the human operator power not used in the tracking task. It is noted that as learning the task improves, the human uses these correlated components more productively.

References

[1] Repperger, D.W. and A.M. Junker, "A Method of Motion Simulator Design Based on Modeling Characteristics of the Human Operator," The Fourteenth Annual Conference on Manual Control, NASA Conference Publication 2060.

[2] Levison, W.H., and G.L. Zacharias, "The Effects of Delayed Motion Cues on Learning and Performance In a Roll-Axis Tracking Task," BBN Report No. 3979, November, 1978.

[3] Levison, W.H. and R.E. Lancraft, "Effects of Simulator Delays on Performance and Learning in a Roll-Axis Tracking Task," The 15th Annual Conference on Manual Control, March, 1979.

[4] Mehra, R.K, "On the Identification of Variances and Adaptive Kalman Filtering," IEEE Transactions on Automatic Control, Vol. AC-15, No. 2, April, 1970, pp. 175-184.

[5] Enstron, K.D. and W.B. Rouse, "Real Time Determination of How a Human Has Allocated His Attention Between Control and Monitoring Tasks," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-7, No. 3, March, 1977.

[6] Repperger, D.W., K.A. Smiles, J.A. Neff, and W.C. Summers, "A Feature Selection Approach in Evaluation of Parameter Changes on The Human Operator Under Thermal Stress," Ergonomics, 1978, Vol. 21, No. 1, pp 35-48.

[7] Gelb, A., "Applied Optimal Estimation," MIT Press, 1974.

[8] Jex, H.R. and A.M. Junker, "Roll Tracking Effects of G-Vector Tilt and Various Types of Motion Washout," The Fourteenth Annual NASA - University Conference on Manual Control, Los Angeles, California, April 25-27, 1978.

Appendix A

It is desired to relate the transfer function in equation (7) into a state formulation in equation (4). The state formulation (4) requires that the pair [A,B] is completely controllable and [A,H] is completely observable in order to use the algorithm in [4]. Equation (7) is written in the form:

$$\frac{\bar{U}(s)}{E(s)} = \frac{d(s+a)}{s^2+s(b+c)+bc} e^{-\tau s} \tag{A.1}$$

in the time domain, a steady state representation of (A.1) would be of the form:

$$\frac{d^2}{dt^2}\bar{U}(t) + \frac{d}{dt}\bar{U}(t)[b+c] + \bar{U}(t)[bc] = ad\, e_\tau(t) + d\left[\frac{de_\tau(t)}{dt}\right] \tag{A.2}$$

where $e_\tau(t) = e(t-\tau)$ which is the stored values of the empirical data.

choose for state variables:

$$x_1(t) = \bar{U}(t) \tag{A.3a}$$

$$x_2(t) = \frac{d}{dt}\bar{U}(t) \tag{A.3b}$$

Then

$$\frac{d}{dt}x_1 = x_2 \tag{A.4a}$$

$$\frac{d}{dt}x_2 = -x_2[b+c] - x_1[bc] + ade_\tau(t) + d\,\dot{e}_\tau(t) \tag{A.4b}$$

or

$$\frac{d}{dt}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} = \begin{bmatrix}0 & 1\\ -a_{21} & -a_{22}\end{bmatrix}\begin{bmatrix}x_1(t)\\ x_2(t)\end{bmatrix} + \begin{bmatrix}0 & 0\\ b_{21} & b_{22}\end{bmatrix}\begin{bmatrix}e_\tau(t)\\ \dot{e}_\tau(t)\end{bmatrix} \tag{A.5}$$

where $a_{21}=bc$, $a_{22}=b+c$, $b_{21}=ad$, and $b_{22}=d$ (A.6)

The complete controllability conditions are checked as follows:

$$\text{rank } [B,AB] = \begin{bmatrix}0 & 0 & b_{21} & b_{22}\\ b_{21} & b_{22} & -a_{22}b_{21} & -a_{22}b_{22}\end{bmatrix} = 2 \tag{A.7}$$

Thus if $b_{21} \neq 0$, $b_{22} \neq 0$, $a_{22} \neq 0$ guarantees the complete controllability conditions.

Referring to (A.1), this requires $a \neq 0$, $d \neq 0$, $bc \neq 0$ which is satisfied by the Bode plot construction.

The complete observability conditions require that:

$$\text{rank } [H^T, A^T H^T] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 2 \tag{A.8}$$

which is always true. Finally it is noted from (A.5) that $a_{21} > 0$ and $a_{22} > 0$ if b and c are chosen from the Bode plot. Hence the A^2 matrix is a stable matrix. From [4] it is ensured that the algorithm will converge to a unique minimum.

Appendix B - Orthogonality Relationships of The Kalman Filter

It is desired to decompose $y(t)=u(t)+n(t)$ into two parts. One part $y_e(t)$ is correlated with the error signal e(t) and the second part $y_o(t)$ is orthogonal to the error signal. To determine these components, the solution to equation (7) can be written:

$$\hat{x}(t) = \int_{t_o}^{t} \phi(t,s)K(s)v(s)ds + \int_{t_o}^{t} \phi(t,s)B\bar{e}(s)\, ds \tag{B.1}$$

where $\phi(t,t_o)$ satisfies:

$$\phi(t_o,t_o) = I \tag{B.2}$$

$$\frac{d}{dt}\phi(t,t_o) = A\,\phi(t,t_o) \tag{B.3}$$

with reference to equation (9), y(t) can be written:

$$y(t)=u(t)+n(t) = H\int_{t_o}^{t} \phi(t,s)K(s)v(s)ds + H\int_{t_o}^{t} \phi(t,s)B\bar{e}(s)ds \tag{B.4}$$

$$+Hv(t)+n(t)$$

and inspection of equation (B.4) yields:

$$\hat{x}_e(t) = y_e(t) = H\int_{t_o}^{t} \phi(t,s)B\bar{e}(s)ds \tag{B.5a}$$

and $\hat{x}_{oe}(t) = y_o(t) = H\int_{t_o}^{t} \phi(t,s)K(s)v(s)ds + H\,v(t) + n(t)$ (B.5b)

The decomposition (B.5a-b) is easily seen to be true because of the result:

$$E\left\{\hat{x}_{oe}(t)\quad \bar{e}(t)\right\} = 0 \qquad \text{follows} \tag{B.6}$$

From the whiteness property of v(t). Hence the expression (B.5b) is the desired orthogonal decomposition and can be computed in real time. In the computer implementation, $\hat{x}_o(t)$ is first computed using (B.5a). To compute $\hat{x}_{oe}$, the relationship $\hat{x}_{oe}{}^{2e} = \hat{x}^2 - \hat{x}_o^2$ is used.

Appendix C - A Measure of Learning In The Tracking Task

With reference to equations (B.5, B.6) and Figure (9), it is desired to compute a measure of learning from variables determined by the Kalman filter. From Figure (9), the error signal E(ω) is known; also available are the stick components $u_a(\omega_1)$ and $u_b(\omega_1)$ which satisfy for each ω_1:

$$u(\omega_1) = u_a(\omega_1) + j u_b(\omega_1) \tag{C.1}$$

and $$E\{u_b(\omega_1)\, e(\omega_1)\} = 0 \tag{C.2}$$

The human's output u(t), however, is processed through the plant H(s). Using phasor notation the plant is written as:

$$H(s) = H(\omega_1) \angle \Theta_2(\omega_1) \tag{C.3}$$

for each frequency ω_1. Writing $u(\omega_1)$ in phasor notation yields:

$$u(\omega_1) = u(\omega_1) \angle \phi_1(\omega_1) \tag{C.4}$$

The component of the error signal due to the hand response becomes (Figure (9)):

$$e_{Stick} = [u(\omega_1) \angle \phi_1(\omega_1)]\; H(\omega_1) \angle \Theta_2(\omega_1) \tag{C.5}$$

$$e_{stick} = uH(\omega_1) \angle [\phi_1(\omega_1) + \Theta_2(\omega_1)] \tag{C.6}$$

With reference to Figure (9) the decomposition of H $u(\omega_1)$ projected on the error axis can be written:

$$u_{e_{correlated}} = \overline{b}-\overline{c} = u_2(\omega_1) \tag{C.7}$$

$$u_{e_{orthogonal}} = \overline{a}+\overline{d} = u_1(\omega_1) \tag{C.8}$$

In terms of e , u_1 and u_2 this can be written:

$$u_{e_{correlated}} = u_a \cos \Theta_2 - u_b \sin \Theta_2 = \overline{b}-\overline{c}=u_2 \tag{C.9}$$

$$u_{e_{orthogonal}} = u_a \sin \Theta_2 + u_b \cos \Theta_2 = \overline{a}+\overline{d}=u_1 \tag{C.10}$$

where Θ_2 is the angle of H(s) at $s=\omega_1$. To determine (C.9) and (C.10)

Spectra are obtained of $u_a(\omega_1)$ and $u_b(\omega_1)$ from the Kalman Filter using the transfer function of the plant [8], the phasor diagram is constructed for a frequency range less than 2 H_z. The components $u_a(\omega_1)$ and $u_b(\omega_1)$ are computed for each ω_1 (knowing the phase angle shift, and magnitude change due to the plant). Summing the power in $u_1(\omega_1)$ over the frequencies gives the stick power not productive in reducing the error signal.

The measure of learning in this task is chosen as:

$$J_1 = \sum_{i=1}^{\bar{n}} \text{Power}\,[u_1(\omega_1)] \Big/ \sum_{i=1}^{\bar{n}} \text{Power}\,[u(\omega_1)] \tag{C.11}$$

Where $\bar{n}$ is the number of frequencies chosen and is a measure of the percent of the stick output not productive in the task of reducing the error signal. This is wasted effort which does not actively help in the tracking task.

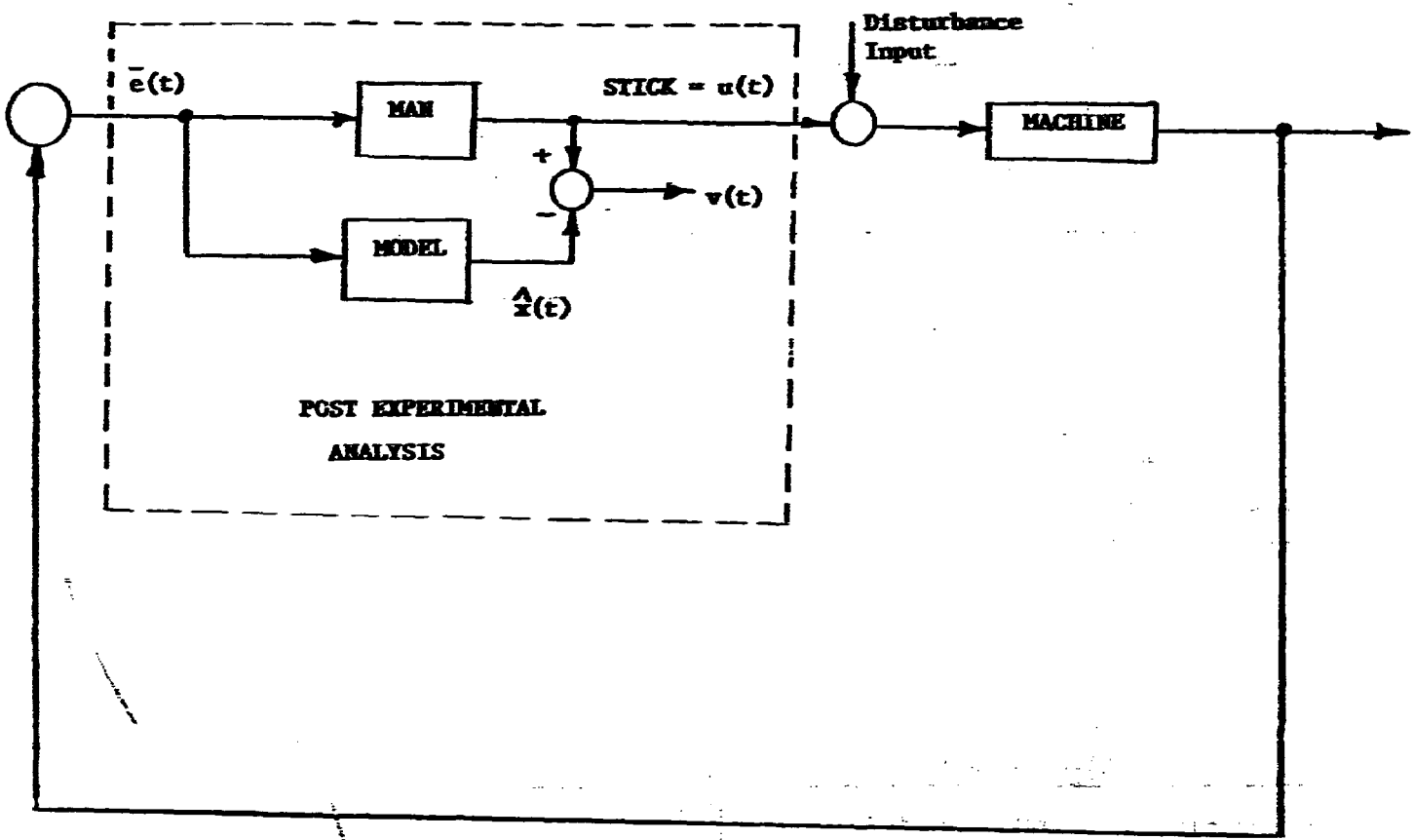


FIGURE (1) - The Man-Machine System

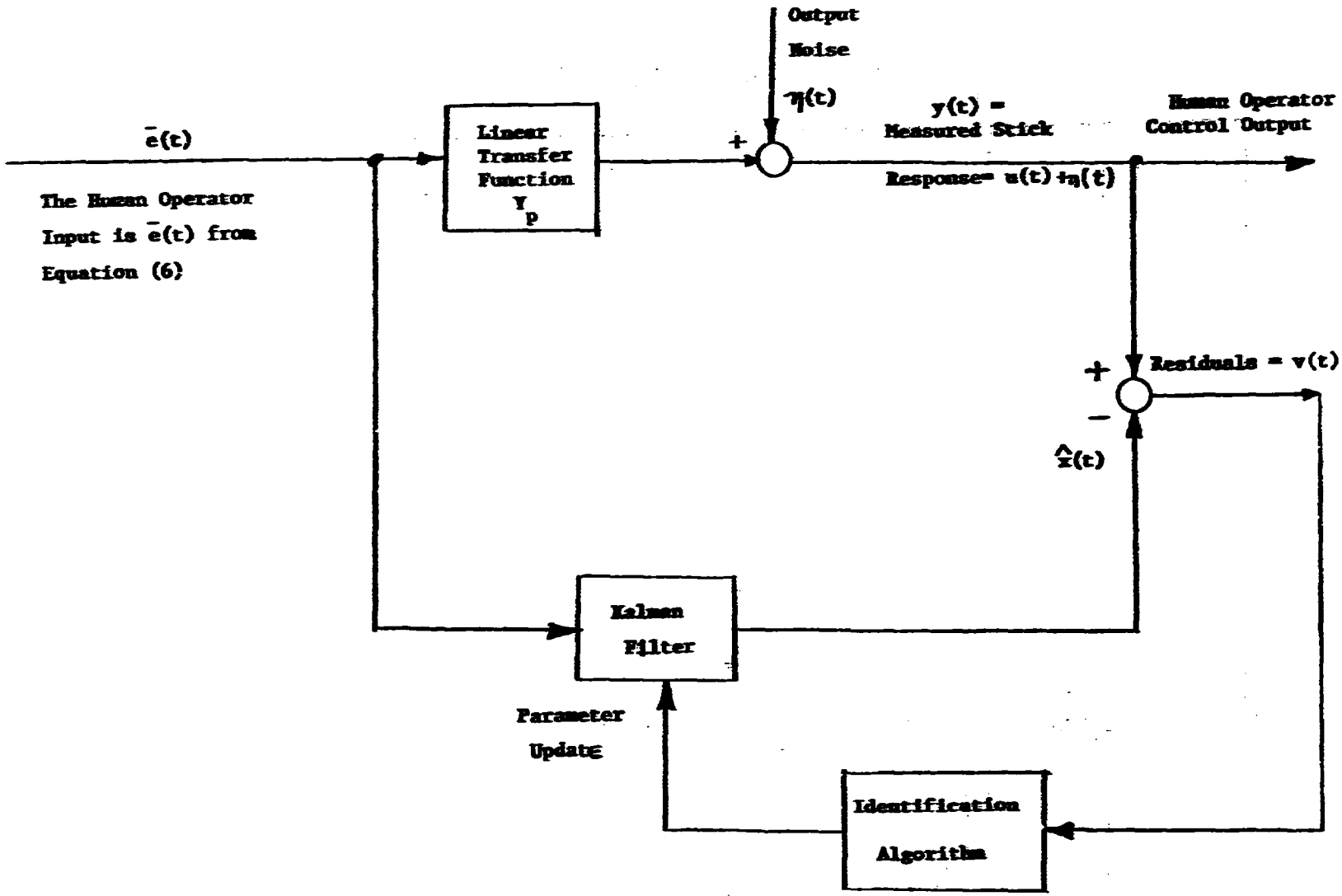


Figure (2) - A Quasi-Linear Human Operator Model and The Kalman Filter

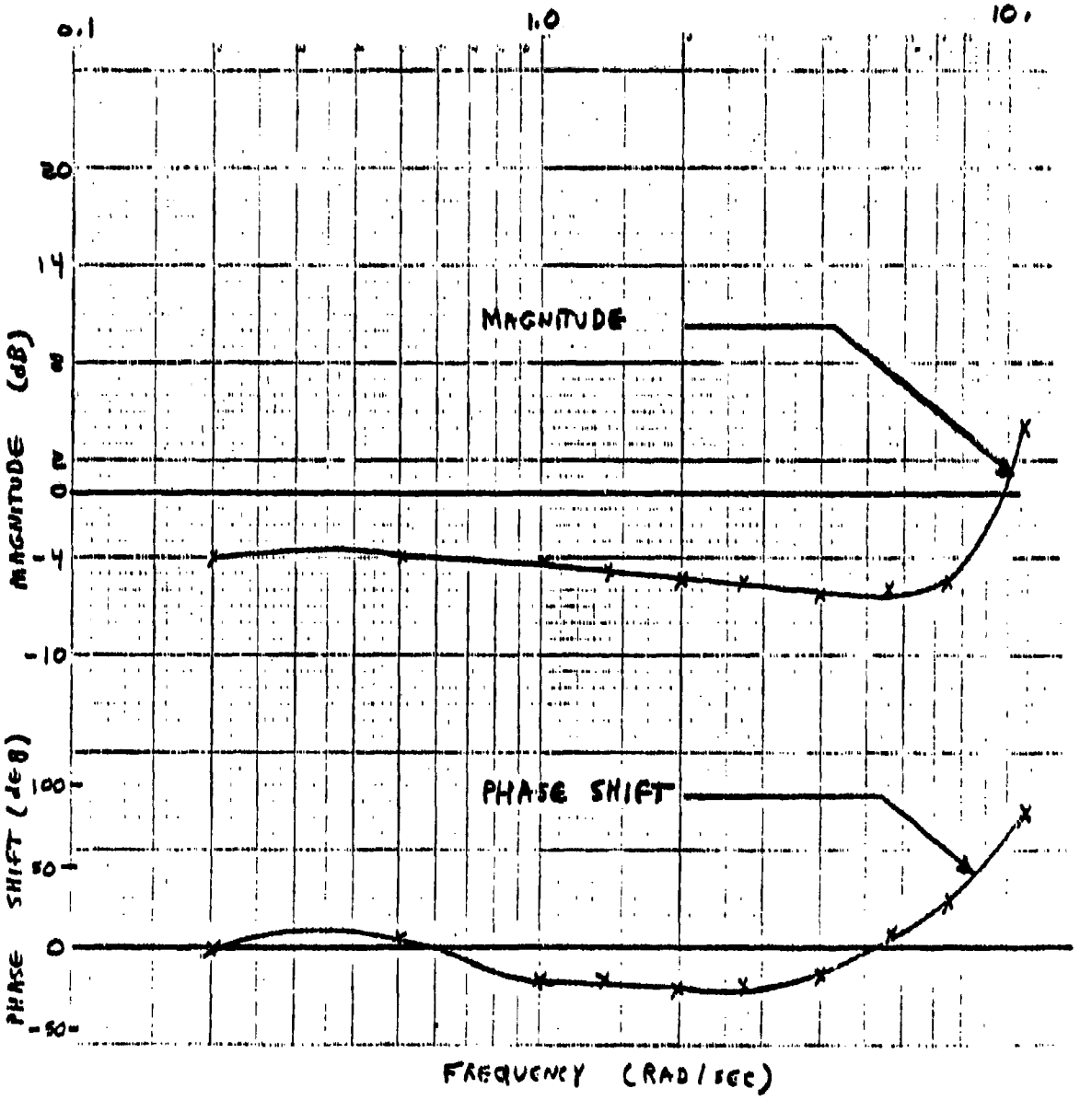


Figure (3) - Magnitude and Phase of The Human Describing Function

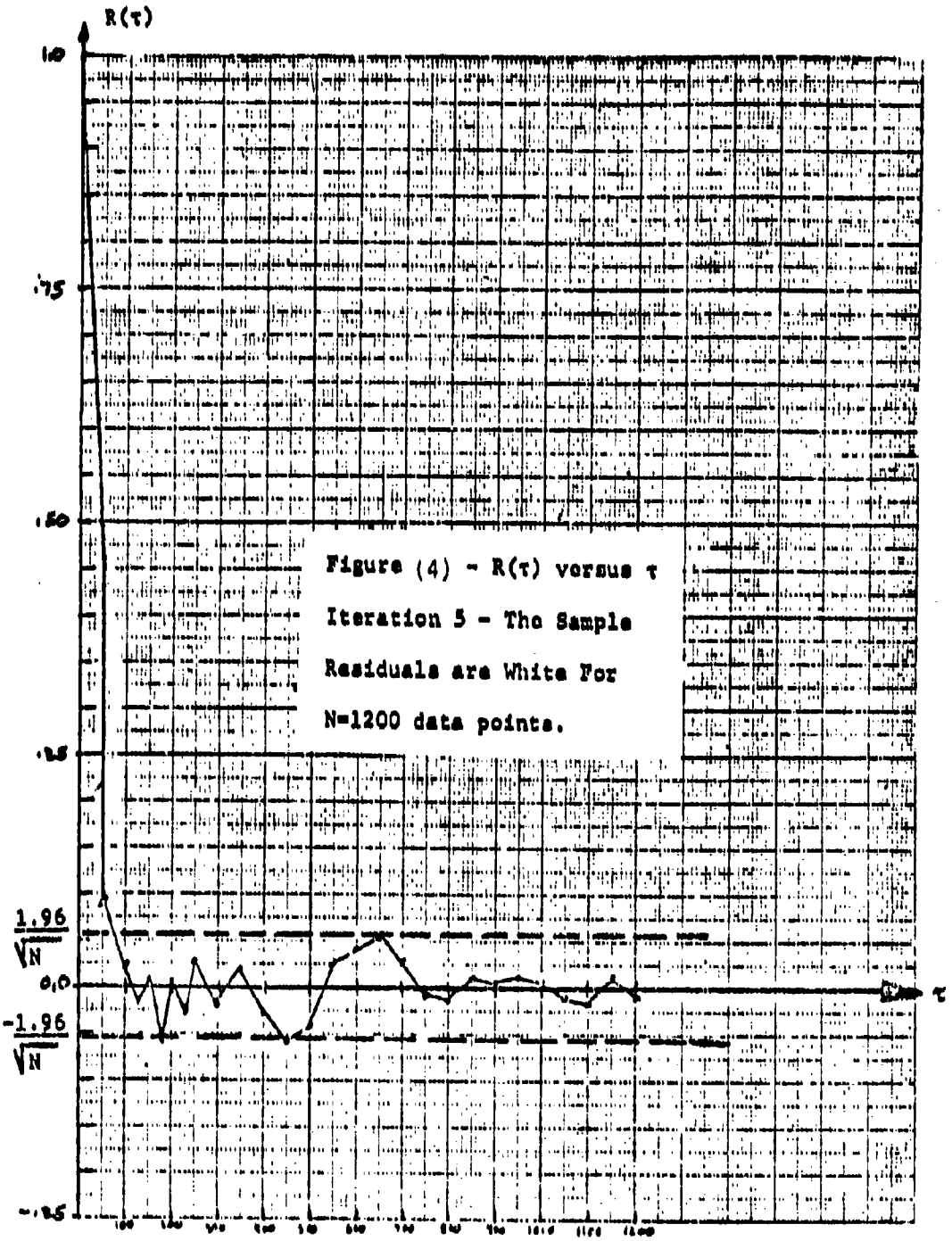


Figure (4) - R(τ) versus τ

Iteration 5 - The Sample

Residuals are White For

N=1200 data points.

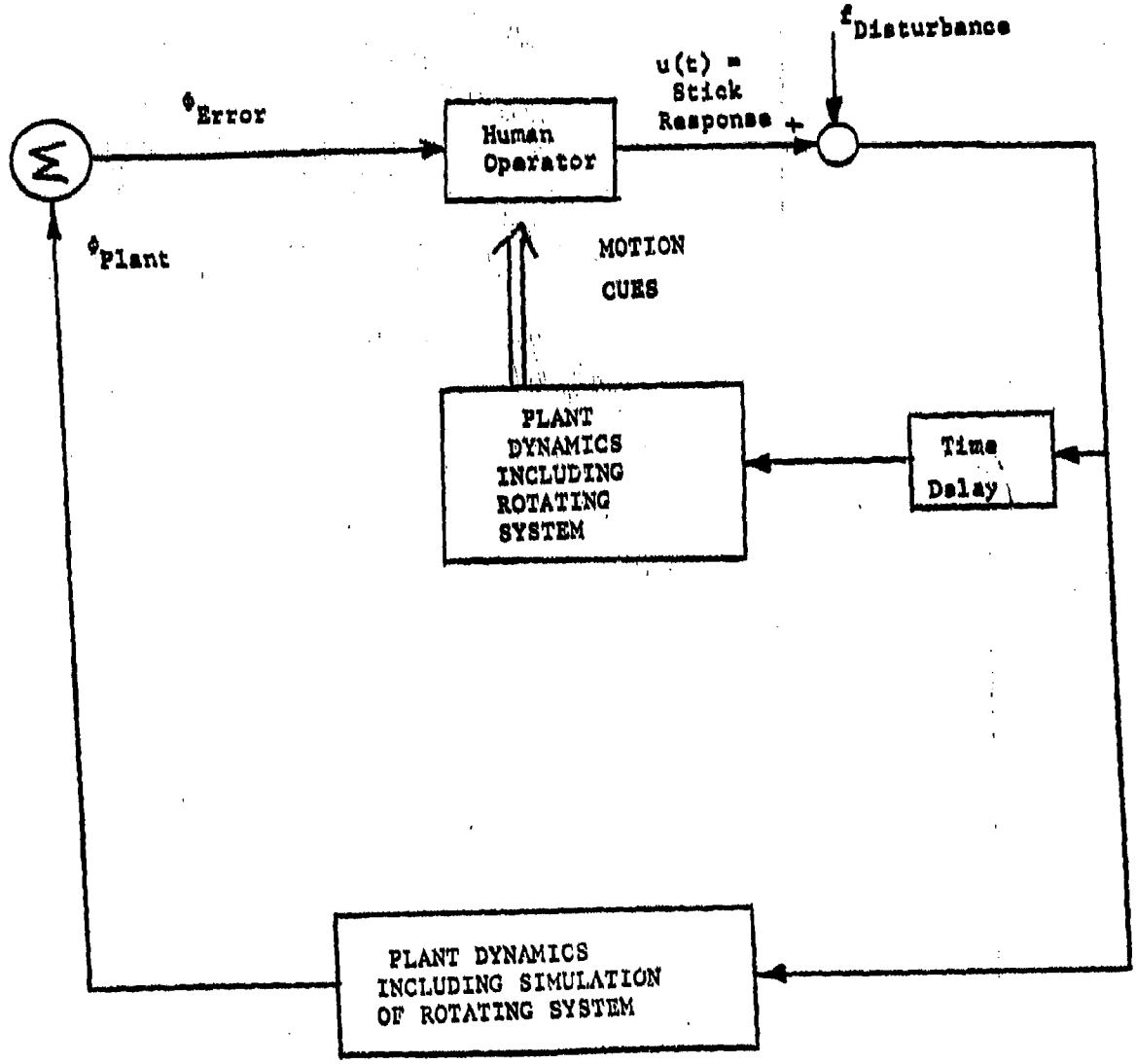
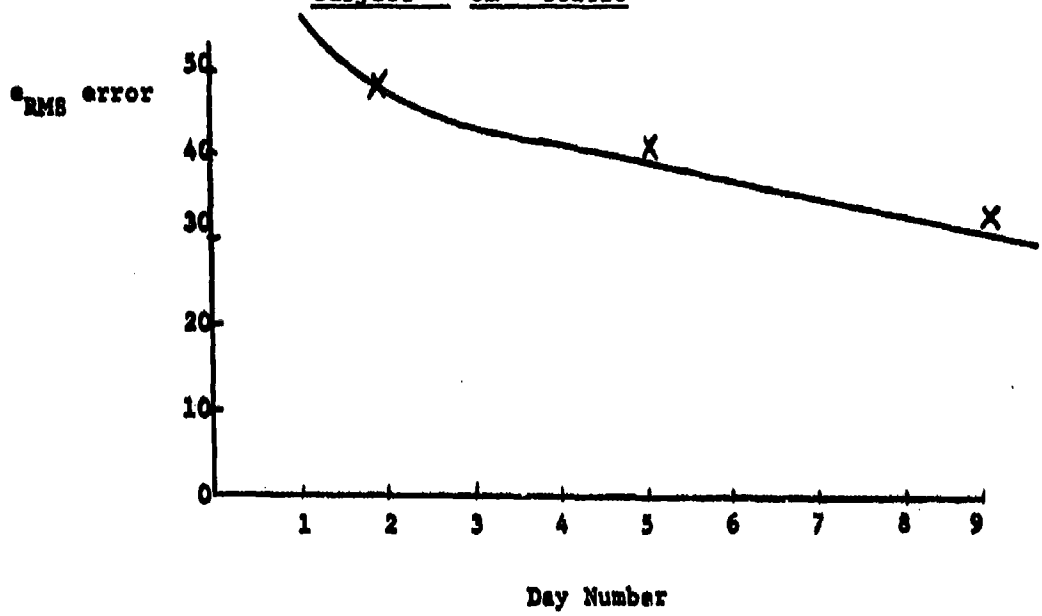


Figure (5) - The Motion Experiment



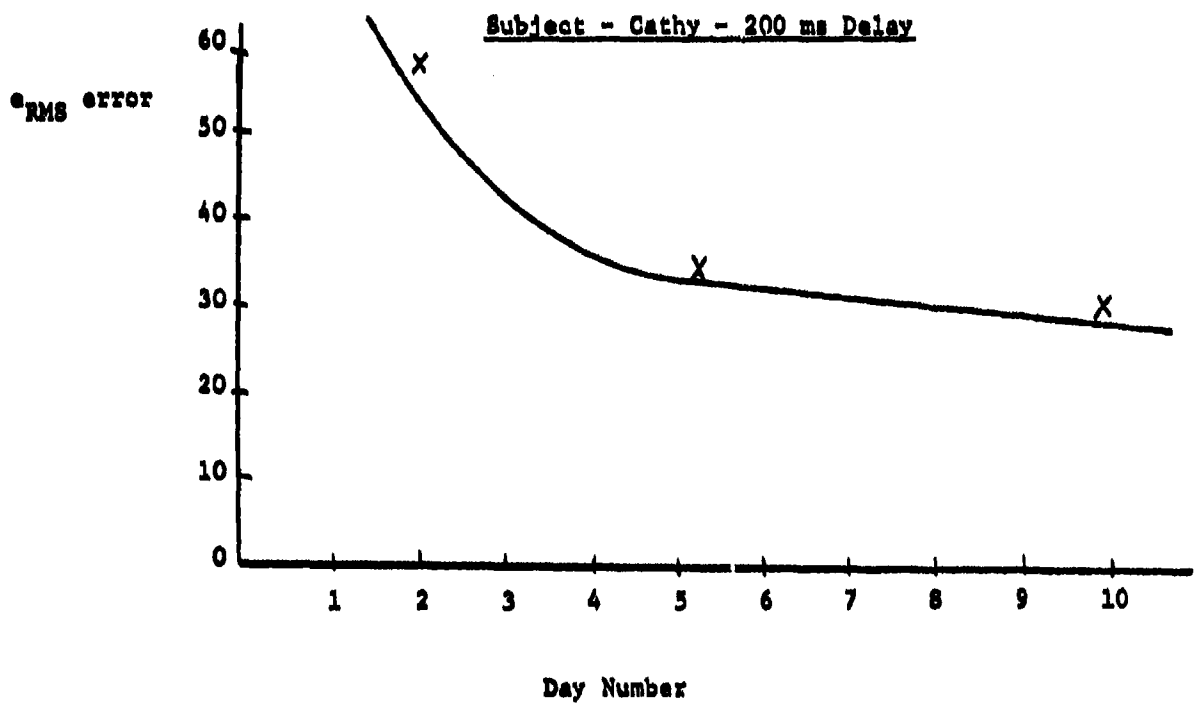


Figure (6) - Data From The Learning Periods

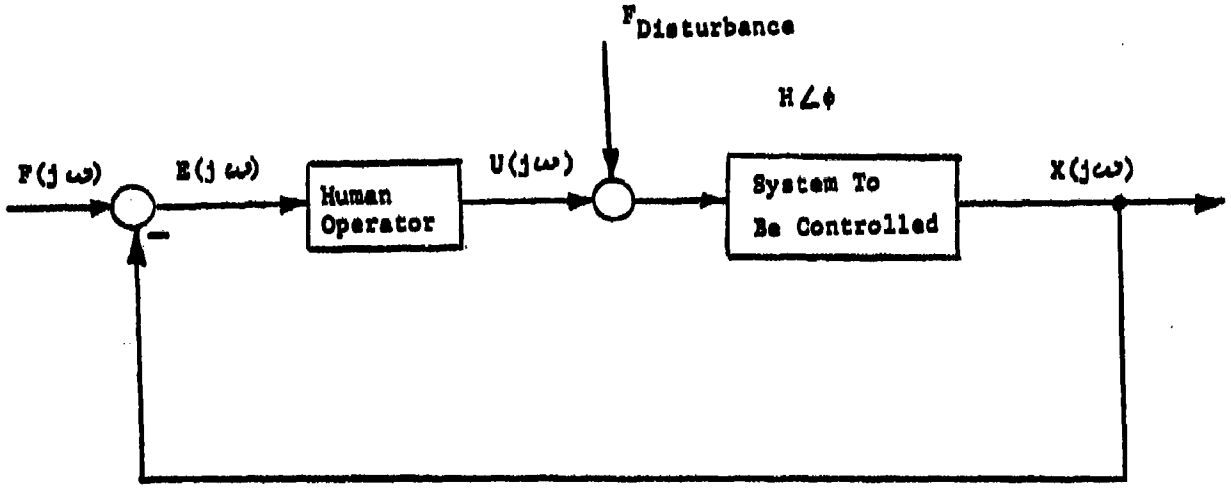


Figure (7) - A Phasor Interpretation

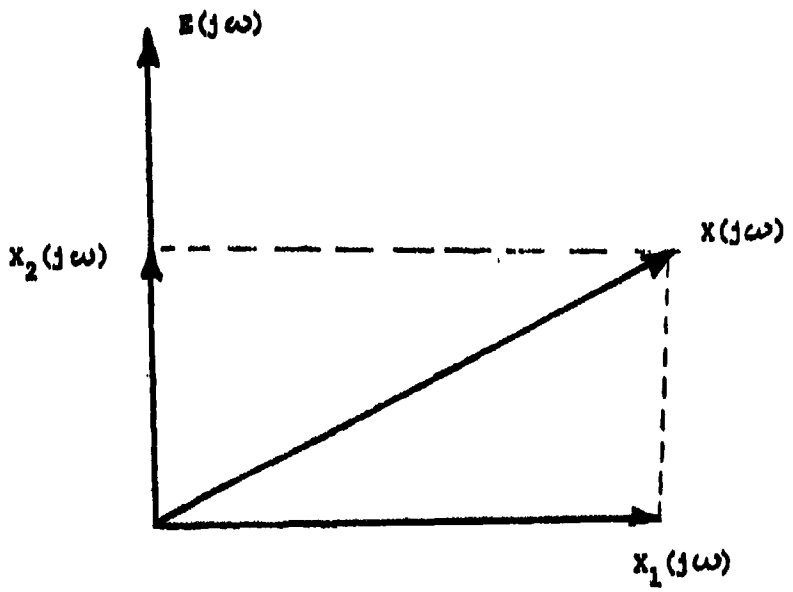
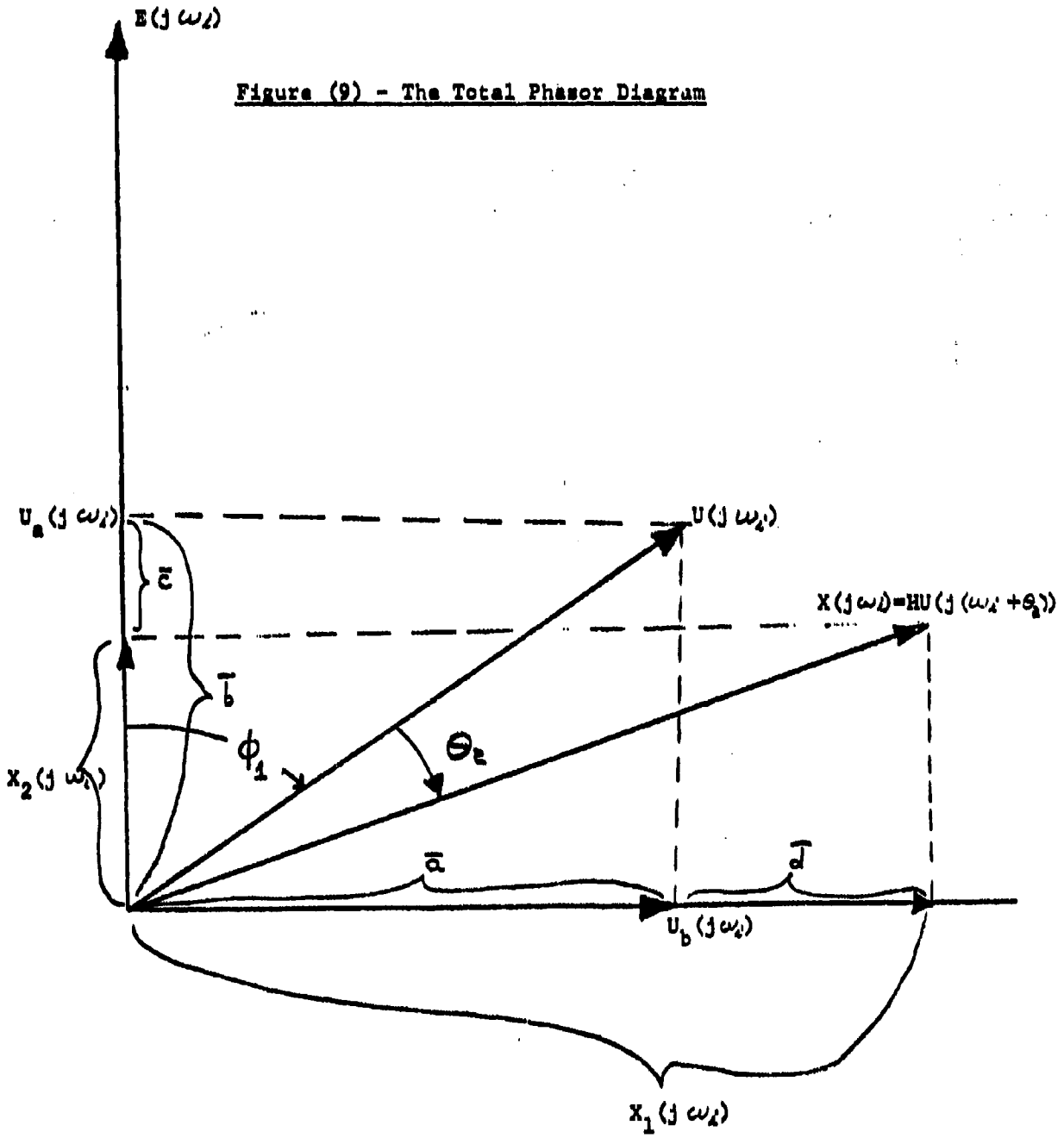


Figure (8) - A Phasor Diagram

Figure (9) - The Total Phasor Diagram



Optimal System Augmentation For a Tracking Task With High-Order Dynamics

by

David K. Schmidt
School of Aeronautics and Astronautics
Purdue University
West Lafayette, Indiana 47907

ABSTRACT

In a previous paper, the theoretical development of an augmentation synthesis methodology was presented. The method utilizes the optimal control model (OCM) of the human (pilot), and applies optimal control theory to synthesize the control laws (vehicle augmentation) compatible with the pilot's objectives and capabilities. To maximize the pilot acceptance, the pilot's objective function is used not only in the development of the pilot model, but also in the augmentation design. The problem then involves the simultaneous solution of two coupled optimal-control problems.

In this paper, the method has been used for the development of integrated fire and flight control laws for augmenting the air-to-air tracking task. This task was previously addressed experimentally and analytically (via pilot modeling) for the longitudinal axis only. In the work to be presented, these previous results for pitch tracking will be reproduced, and then a family of control laws developed, tracking improvements predicted, and system-dynamic augmentation discussed. The plant dynamics will not only include the vehicle rigid body dynamics, but the dynamics of several lead-computing sights. The lead-computing optical sight (LCOS) as well as a perfect director sight at two tracking ranges and two levels of rms target activity are considered.

An interesting result is the relationship between the desired vehicle dynamics (e.g., short period) and the sight dynamics, demonstrating the importance of considering all the dynamic elements in developing handling qualities specifications, for example.

AN ECLECTIC REFORMULATION OF FLYING QUALITIES

E. D. Onstott
W. H. Faulkner

Northrop Corporation
Aircraft Group
Hawthorne, California

ABSTRACT

The development of high authority and non-standard control configurations has led to aircraft designs that cannot be well correlated with the existing flying qualities data base. For the subject of flying qualities to be responsive to current needs in aircraft design and procurement, an eclectic reformulation of the subject is required. This paper presents one approach that allows much greater freedom in problem formulation and suggests ways to free the subject area from the domination of specialized methodologies.

INTRODUCTION

The authors have recently completed a study for the USAF Flight Dynamics Laboratory which resulted in the documentation of a comprehensive account of one particular approach to the prediction, evaluation, and specification of flying qualities. This time history simulation technique was not intended to replace or render obsolete other flying qualities approaches; rather it was initially developed to analyze problems that were not amenable to study by linear or time-invariant means. Since the publication of the contract report, AFFDL-TR-78-3, the authors have received a number of comments that indicate two widespread difficulties:

- The subject of flying qualities is often defined by its practitioners in terms of existing analytical and test methods.
- It is difficult for new approaches to the subject to be properly understood in relation to the existing technology.

It is the object of this paper to indicate how the subject of flying qualities inherently contains the means to overcome the limitations of established methodologies

so that a fully eclectic capability will be available for the study of aircraft now being designed. Two things will be presented:

- A discussion of how flying qualities can be defined with respect to its functions: evaluation, specification, and prediction.
- An indication of how the definitions and subject requirements of flying qualities can be interpreted in terms of specific methods of testing, specification, analysis, and prediction.

Conventional flying qualities practice implicitly depends on the almost universal dynamic similarities of fixed-wing aircraft. This allows a direct comparison of airframe dynamics that are easily calculated and correlated with operational experience. However, the recent development of high authority control systems which tend to obscure the basic airframe dynamics implies that the conventional aircraft comparisons are no longer sufficient as the basis of flying qualities analysis.

Furthermore, the domain of flight controls and flying qualities is becoming considerably broader as a result of:

- Integration of control, propulsion, weapons, and navigation systems.
- Ability to decouple and regroup the aircraft dynamic modes.
- Advanced head-up and CRT displays.
- New controller concepts.

For these reasons, experience based on aircraft comparisons cannot constitute a complete basis of flying qualities work in future applications. Thus, to be responsive to the needs of aircraft design and procurement, the subject of flying qualities must be reformulated in accordance with the following postulate:

All flying qualities evaluation, specification, and prediction concerns must be expressed in terms of a particular airplane and how well it can meet intended design and procurement objectives.

The practice of flying qualities has started adopting this basic postulate in the following areas:

- Flight test and simulation – development of standardized test maneuvers, Reference 1.
- Specification – proposed USAF Prime Standard and Handbook, Reference 2.
- Analysis and Prediction Methods – CCV studies, target tracking analysis, discrete maneuver analysis, Reference 3.

These trends reflect the importance of asking the right questions in terms of what the pilot can make the airplane do, independent of established analysis methodologies.

In order to see how the basic postulate can be followed, it will be useful to examine the resources of flying qualities as a subject and how they can be used for the principal applications of flying qualities evaluation, specification, and prediction for each area of concern.

DEFINITION OF FLYING QUALITIES

Historically, the subject of flying qualities has developed by utilizing available analytical and experimental methods. These constitute the subject's four basic resources:

- Historical data and data correlations.
- Flight test procedures and capabilities.
- Flight simulation methods, both in-flight and ground-based.
- Mathematical modeling and analysis techniques.

For example, the description of aircraft dynamics, both open and closed loop, has followed the trends of control theory. Frequency response descriptions were first employed, but with the development of root locus methods, eigenvalue descriptions became widely used. At present, state variable and optimal control methods are becoming prominent along with time history simulations. At each stage of this development, there were increased capabilities to correlate aircraft dynamics descriptors, and to define further aspects of flying qualities consideration in terms of available dynamic models. In this way, the available or accepted methodologies determined what aspects of piloted flight were appropriate for study.

If flying qualities is to become fully responsive to advanced aircraft design and procurement, this process must be reversed. By investigating definitions and objectives of the subject, guidelines for the supporting methodologies can be established; the following discussion will do just this.

In developing the subject independent of methods, it should be kept in mind that there are a large number of dynamic aspects of piloted flight that are largely independent in description. This diversity includes, for example,

open loop response	display and cockpit layout
precision tracking	flight safety
PIO	formation flight
integrated fire-flight control	feel system characteristics

and the investigation of a particular aircraft must include all such relevant items of concern.

In terms of the above considerations, the subject of flying qualities can be defined as the discipline of investigating all relevant areas of concern by means of the above resources for the following purposes:

- Judging the performance of a specific airplane.
- Aircraft procurement and design specification.
- Aircraft performance and flying qualities prediction for use in:
  - Aircraft design and development.
  - Aircraft improvement and modification.

AIRCRAFT EVALUATION

Before judgments can be made concerning piloted performance, data must be obtained from, or assigned to, the airplane under consideration. This process of obtaining descriptive piloted performance data is called aircraft evaluation and data obtained can be classified as follows:

- Objective - numerical measures obtained through instrumentation.
- Subjective - pilot statements.
- Analytical - behavior of mathematical aircraft dynamic models.

The totality of these evaluation data for a given airplane is called its flying qualities, while the totality of the subjective evaluation data is often referred to as the handling qualities.

Objective Evaluation Data

Objective data can be routinely obtained for all aspects of aircraft performance. Consequently, there is no restriction in applying the basic postulate of flying qualities with respect to aircraft evaluation using objective data.

Subjective Evaluation Data

The pilot's subjective evaluation consists of how well he thinks the airplane did or could do, and how much "workload" was involved, supported by diagnostic comments about good or deficient airplane characteristics. This information can be obtained

from the pilot for each area of concern, or as a summary of particular aircraft flight or mission phases. This subjective evaluation is always available in conjunction with any flight test item that leads to objective evaluations.

Although there are two possibly independent aspects in the pilot's evaluation, performance and workload, there has been a highly successful and almost universal method of reducing the subjective evaluation to a scalar quantity, namely the pilot opinion rating, Reference 4. Since pilot acceptance of an airplane is of great importance along with acceptable objective performance, the use of pilot opinion rating methods, particularly the Cooper Harper rating scale, should be continued.

Unfortunately, there has been no standard method for supporting the rating by diagnostic comments. This means that it is seldom possible to understand the blend or compromise of performance and "workload" that a pilot has used in giving his rating judgement, even though a decision tree of performance and compensation descriptors is explicitly provided in the Cooper Harper rating method. This is further confused by diverse assumptions adopted by many flying qualities analysts that workload is:

- compensation
- identified pilot parameters
- physical exertion against controller force gradients
- reserve attention
- time estimation
- total angular rate
- pilot - aircraft payoff functional

Depending on the flight task evaluated, each of the above interpretations of workload may be the most meaningful or influential on the pilot rating. In the case of an airplane with a high workload, it is important to understand the exact dynamic nature of the problem before corrections to aerodynamics or control modifications can be undertaken. For these reasons, it is recommended that in support of a Cooper Harper pilot rating, the pilot state what aspects of workload seem to dominate. This should be included in a standardized diagnostic questionnaire. This recommendation implies no change in the Cooper Harper rating scale or methods; it applies exclusively to the way and precision with which the scale is employed.

Analytical Evaluation Data

Analysis as a means of aircraft evaluation must be carefully understood. As an evaluation, any performance measure obtained by analysis must be regarded as equivalent to what would be measured using the actual airplane. Thus a valid analytical evaluation has the same evaluation status as actual flight test data. Examples allowed by Air Force specification practice include dutch roll and short period eigenvalues calculated from determined airframe aerodynamics, and dutch roll mode amplitude and phase measures also calculated from aerodynamic data. In this way, the calculated eigenvalues, for example, are used to determine acceptable or unacceptable performance according to Reference 5, Military Specification, Flying Qualities of Piloted Airplanes, MIL-F-8785B.

For this reason, analysis methods must be far more reliable than prediction methods, whose main purpose is to guide early design and development. In order to assure the required reliability, the following conditions on the analytical methods should be met whenever possible:

- All calculated quantities should be potentially measurable or reducible from flight test data.
- Verifications by flight test data should be obtained for representative flight conditions.
- All data used in the calculation must be either obtained from, or verified by, flight testing.
- All aircraft, control, and aerodynamic models, regardless of formulation in the time domain, s-plane, or state space, must be sufficiently general to include all relevant dynamic and kinematic effects.

It was stated in the Introduction that the basic postulate of flying qualities implied that questions concerning a particular airplane should be phrased in terms of its particular design and procurement objectives. This means that to be most meaningful, the following three questions must be resolved prior to flight testing and analysis:

- What test maneuvers are to be flown?
- Which items are to be evaluated by the same test maneuver?
- For what purposes are the objective and subjective data to be used?

AIRCRAFT SPECIFICATION

Specification consists of criteria by which judgments are assigned to aircraft evaluation data. These come about in the following way: Procurement objectives state what the intended airplane must be able to do, while specification criteria, whether developed by the procuring agency or the airplane designer, express how well the airplane should perform in terms of aircraft evaluation and pilot acceptance as discussed above. The overall objectives of satisfying the specification criteria are:

- guaranteed aircraft capability
- guaranteed pilot acceptance

Procurement criteria, such as MIL-F-8785B, have evolved by identifying correlations among conventional aircraft between performance measures and acceptable pilot ratings. In this way, the evolved criteria have dictated the performance measures to be evaluated along with associated flight test and analysis methods. Since this approach requires comparison of many similar aircraft, specification methods must be substantially augmented for current and future specification of unconventional new designs and mission roles.

On the other hand, if evaluation measures have been comprehensively developed for a particular airplane in accordance with the basic postulate, then all that is required are decisions on how well the airplane should perform on each objective and subjective evaluation item. In this way, the most meaningful evaluation of a particular airplane dictates an appropriate specification with respect to the procurement and design objectives.

Introduction by the United States Air Force of the USAF Prime Standard and Handbook to supersede MIL-F-8785B, Reference 2, will achieve the above objective supported by the data base of the current specification and its background and user guide, Reference 6. The large number of independent evaluation items will require a widely diverse supporting technology. This must be well represented in the new documents which will indicate many ways to approach evaluation and specification.

It should be noted that the flying qualities analysis of future aircraft performed in accordance with the principles outlined above, will proceed by selecting the most appropriate and comprehensive objective and subjective evaluation items, placing specification requirements on them, and only then, selecting analysis and prediction methods to support the design and development of the airplane. It was pointed out above that prediction and analysis methods are used in entirely different ways; this will next be considered in more detail.

PERFORMANCE AND FLYING QUALITIES PREDICTION

Prediction of flying qualities consists of developing and exercising mathematical models of open loop and closed loop aircraft response. The objectives of these analyses are to:

- Predict compliance of evaluation items with specification criteria.
- Predict probable performance, pilot acceptance, and dynamic characteristics.
- Predict performance tradeoffs among design parameters.

Inasmuch as future aircraft evaluations and specification criteria will be generated according to design and procurement objectives, many new kinds of prediction methods must be developed and validated. These methods must predict both objective and subjective evaluation data for all areas of flying qualities concern. In this way, the selection of evaluation parameters and specification criteria will lead to selection of the appropriate pilot - aircraft models and prediction techniques. This resulting prediction methodology will be useful in the following ways:

- To guide preliminary aerodynamic and control design.
- To guide final design during aircraft development.
- To identify, understand, and eliminate flying qualities deficiencies.
- To assist in demonstrating compliance with procurement and design objectives.
- To assist in interpretation of pilot ratings and comments.
- To search for and identify unrecognized but relevant flying qualities phenomena.

The selection of appropriate prediction methods depends upon the representation of a particular item of flying qualities concern. This representation will always consist of three separate parts. They are:

- Task Model. This is a mathematical description of a sufficiently representative flight test item.
- Aircraft or Pilot - Aircraft Model. These models represent the dynamics of the airplane or the closed loop piloted response during the performance of the task as represented by the task model.

- Evaluation Model. The evaluation model should include all objective evaluation items that would be obtained during the corresponding flight test. The evaluation must also include data which can be shown to correlate well with subjective pilot ratings and comments.

Once these model components have been chosen, the particular techniques for calculating the evaluation quantites can then be selected or developed.

Task Model

Task model selection in addition to being required for prediction methods, is also an important aspect of flight test programs. For example, Twisdale, Neal and Smith, and Meeker and Hall, References 1, 7, 8, have performed extensive flight test and in-flight simulation studies using target tracking during wind-up turns, attitude tracking of random commands, and step attitude tracking. These task models were selected to be representative of combat tracking, and have proved to be good predictors of aircraft operational experience. The success of these models makes them likely candidates for use in the prediction of combat tracking by means of closed loop pilot - aircraft modeling methods.

Aircraft Model

Aircraft model selection depends on the task model in the following ways:

- Flight condition. High angle of attack or sideslip angles may require nonlinear aerodynamic data.
- Maneuvering required. Large angular excursions and high angular rates may require nonlinear coupled equations to represent adequately the performance of the task model.
- Control system characteristics. If the performance of the task model results in limiting of rates or control surface excursions, these effects must be included in the model. Other dynamic effects such as control augmentation saturation must also be included.

It should be noted that the selection of an appropriate model does not necessarily imply the selection of a computational method, but determines only the necessary dynamic capabilities a computational context for the problem must possess.

The principal uses for the aircraft model are:

- Open loop aircraft analysis.
- Closed loop aircraft analysis.
- Use of the aircraft model for manned flight simulation.

Flight simulation is a reliable method for predicting both objectives and subjective evaluation data. However, the reliability of these predictions depends heavily on such factors as airplane model fidelity, visual and motion cue fidelity, and computational efficiency in both ground-based and in-flight simulations.

Pilot — Aircraft Model

Pilot - aircraft model selection depends on both the task model and the relevant evaluation items. The aircraft part of the pilot - aircraft model should be chosen as above, and the pilot model must be expressed in whatever computational context this requires. Unfortunately, the use of pilot models in flying qualities analysis has remained controversial and non-standardized for the following reasons:

- The pilot model has usually dictated a strictly linear and time-invariant problem formulation.
- The pilot model is often not well defined in terms of what model components are to be used, what dynamic limitations apply, and what adjustment or optimization rules are to be followed.
- The calculation methods are often obscure and the data from the model difficult to compare with flight test or simulation results.

Recent extensions, Reference 3, in pilot model theory have eliminated the linear and time-invariant restrictions, and models can now be chosen to fit any computational context required by the problem formulation. More precisely, all currently used models are special cases of the following definition which will be adhered to in the subsequent analysis:

> Definition: A pilot model is a rule that assigns a dynamical description of a pilot's activity during a given task along with a method for using the model to predict evaluation data. This dynamical description is subject to human limitations that include:
>
> - transport time delay
> - human visual resolution and motion perception thresholds
> - limited motor information output channels
> - neuromuscular dynamics effects
>
> For precision control tasks, the model is adjusted to produce optimum performance with respect to the evaluation parameters. This adjustment may incorporate time-varying compensation, attention allocation, discrete control inputs, and other control strategies that can be identified in human pilot activity.

The selection of a particular set of pilot model characteristics should be made exclusively on the basis of relevant pilot activity for each flying qualities evaluation task item. This implies that the model must be developed independent of computational method, and that various models and computational methods might be required in the study of any given airplane.

Evaluation Model

The evaluation model consists of a set of performance quantities to be calculated from the task, aircraft, and pilot - aircraft models, along with the methods to be used and the interpretation procedures to be applied. All objective evaluation items established through design and procurement requirements for flight testing, can be predicted by means of a properly selected model or manned flight simulation.

It often occurs during design and development that performance quantities are identified that relate to flying qualities in previously unrecognized ways. Such quantities are often related to control or weapons system behavior, and can best be studied using a combination of prediction methods based on analytic models and manned flight simulation. Subjective data obtained from manned flight simulation should be obtained in the same manner as in actual flight test. In interpreting this data, it should be kept in mind that activity of a simulation pilot is, in fact, only a representation of what a pilot would do in the actual aircraft.

Correlations of Pilot Ratings

The prediction of subjective evaluation is performed by postulating correlations of pilot ratings and comments with open and closed loop airplane and pilot model parameters, and performance evaluations.

Correlation of pilot ratings and comments with open loop airplane characteristics is the basis of many items in MIL-F-8785B. For sufficiently conventional aircraft dynamics, short period and dutch roll eigenvalues correlate with pilot ratings. For this reason, acceptable performance is judged when the corresponding eigenvalues are within the specified bounds of frequency and damping. Other open loop dynamics correlate with pilot comments and pilot ratings. Dutch roll amplitude and phasing, for example, are unpleasant to the pilot in cases identified by MIL-F-8785B. Unfortunately, such correlations obtain only for aircraft that have similar dynamic response modes.

Correlations of pilot ratings and comments with closed loop model and performance parameters are applicable to a larger class of aircraft than the open loop correlations. In this case, the correlations are in terms of how well the precision task model is carried out along with some knowledge of predicted pilot activity as reflected in pilot model parameters. There are three basically different methods currently under development or in use.

- Payoff Functionals. These are usually computed by means of optimal control theory and give a weighted blend of output statistics and pilot activity. The values of the optimized functional are correlated with pilot ratings.

- Pilot Rating Formulas. These models postulate that pilot workload is equivalent to pilot compensation so that pilot ratings become functions of pilot model parameters and output statistics weighted in a suitable manner.

- Multi-Parameter Performance Correlations. These methods recognize that many tasks consist of several mutually compromising objectives that the pilot must trade off against one another. By correlating ratings and comments with these tradeoff elements, regions of output statistics that correlate with ratings can be demonstrated without assuming functional definitions of pilot ratings or pilot workload.

Analysis and Prediction Methods

It is extremely important to understand the distinction between analysis methods and prediction methods. Inasmuch as both are used to generate data that is compared against specification criteria, it is natural that confusion exists about the roles of analysis and prediction.

As stated above, compliance with a procurement specification item must be on the basis of aircraft evaluation data obtained from flight test. This data may be transformable in certain ways, such as calculating short period eigenvalues from flight test aerodynamic data, that involve minimal assumptions on the model. Such highly reliable and validated analysis methods simply transform flight test data into a different form.

Prediction methods, on the other hand, are developed for use in design and development as a guide when flight test data are not available. It may well be the case that prediction methods have been instrumental in defining test items and corresponding criteria, but this in no way implies that data generated by these methods are suitable for demonstrating compliance and hence justifying procurement. This is especially true of pilot - aircraft methods, and pilot rating prediction in particular.

The above discussion of the definition and objectives of flying qualities emphasized two principles:

- Formulating flying qualities questions in terms of a particular aircraft and its procurement and design objectives.
- Defining required flying qualities models of task, aircraft or pilot-aircraft, and evaluation, independent of computational algorithms.

These principles implicitly assume that for whatever problem formulation may be developed, suitable means for flying qualities evaluation, specification, and prediction are available. The goal now is to show how flying qualities defined by these methods can be interpreted in terms of specific numerical quantities and computational techniques.

Flying qualities is defined above in terms of its three areas of application: evaluation, specification, and prediction. Moreover, it was shown that the implementation of the basic principles of flying qualities leads to a choice of evaluation items independent of methodology from which specification items can be derived along with appropriate prediction techniques for aircraft design and development. For these reasons, the interpretations of evaluation, specification and prediction will be discussed in that order.

EVALUATION METHODS

Evaluation was defined to be the process of assigning three kinds of data to a specific airplane:

- objective - numerical measures through instrumentation
- subjective - pilot comments and ratings
- analytical - behavior of mathematical aircraft or pilot-aircraft models.

Once these evaluation data are obtained, they can be compared against specification criteria and a judgment of flying qualities goodness can then be made.

The principal source of both objective and subjective evaluation data is, of course, flight testing and eventually operational experience. Since evaluation data is mainly used to compare against procurement criteria, operational data is usually not available.

This means that the flight test programs must be sufficiently representative of the operational experience to give a realistic description of airplane capability.

Flight test procedures have been developed and refined in conjunction with aircraft development and operational aircraft experience, and effective test maneuvers and instrumentation have been developed. This experience has been exclusively concerned with defining how the airplane is to be tested, what tasks are to be flown, what measurements best reflect the resulting performance, and how the pilots are to be trained, introduced to the tasks, and questioned. The emphasis of flight test methodology on obtaining evaluation data (again contrasted to comparing the evaluation data against criteria and making judgments) in terms of the behavior of a specific airplane is in complete accordance with the basic principles of flying qualities as presented above. For this reason, the concerns of flying qualities should be dictated by current flight test practice and trends, and not by convenient computer codes.

The main problem in flight testing is the selection of flight test items. It is important to keep in mind that these items are chosen not as typical operational flight tasks, but as sources of evaluation data that will be as discriminating as possible when used for comparison against design and procurement criteria. This is made most clear in Reference 1 by Thomas Twisdale of the Air Force Flight Test Center:

> "It is very important not to confuse tracking test techniques with the operational tracking and gun firing techniques associated with an actual combat encounter. Tracking test techniques are a powerful tool for identifying and defining handling qualities deficiencies and optimizing flight control systems. These techniques were specifically developed to elicit engineering data which may be used to improve the handling characteristics of the airplane. In this respect it is certainly expected that the results of tracking test techniques (a better handling airplane) will favorably impact the operational pilot's ability to control his aircraft during combat encounters. But it would be a mistake to assume that the data gathered using these techniques directly reflect such overall mission effectiveness parameters as the likelihood of a kill. The overall combat effectiveness of the airplane is a function of many considerations. Tracking test techniques provide a measure of that portion of mission effectiveness which is related to the pilot's ability to precisely control the aircraft attitude."

The tracking test techniques to which Twisdale refers, concern tracking a target aircraft during smooth wind-up or constant angle-of-attack turns. Data of this kind has proved highly useful in a number of aircraft development programs, and this approach to flight testing is firmly established.

Unfortunately, these developments in aircraft evaluation methods have not been matched in flying qualities specification and prediction techniques. Presently there are no criteria in MIL-F-8785B for target tracking performance based on standardized target tracking tasks. This is in part a result of the previous lack of suitable established prediction methods necessary for such criteria to be useful during preliminary design and aircraft development.

Another tracking technique that is gaining prominence in flight testing is in recognition that a major concern of the pilot is his ability to carry out discrete aircraft flight path or attitude corrections quickly without resulting difficulties such as overshoot, residual oscillations, or unfavorable mode coupling. Discrete-error tracking tasks are now frequently used in flight test and flight simulation programs. One of the first and most important flying qualities flight test programs to incorporate these tasks was the study performed in 1970 by T. Peter Neal and Rogers E. Smith using the NT-33 variable stability in-flight simulator, Reference 7. It is useful to consider the pilot evaluation tasks that were used in this study.

An examination of these task items reveals that other than the IFR continuous random tracking task, all flight task items were of a discrete nature, either to make specific control corrections, or to perform open loop maneuvers. In Section 6.2 of Reference 7 titled "The Pilot's View of Good Tracking Performance," Neal and Smith comment:

> "The first step in the analysis is to identify the performance which the pilot is trying to achieve when he "adapts" to an airplane configuration. The pilot comments indicate quite clearly that he wants to acquire the target quickly and predictably, with a minimum of overshoot and oscillation. The question that remains is how to translate this observation into mathematical terms."

This "translation" properly belongs to the subjects of flying qualities specification and prediction and has been addressed in References 3 and 8.

It is interesting to contrast the flight test methods of tracking target aircraft in wind-up turns with the flight test evaluation items based on discrete control corrections and maneuvers. They are in no way mutually exclusive; rather they simply reflect evaluation emphasis on objective items (tracking statistics) or subjective items (pilot ratings and comments).

SPECIFICATION METHODS

The extensive integration of controls, weapons, navigation, and avionics systems in current aircraft designs implies that flying qualities specification items of aircraft performance for all mission tasks cannot be independent. For this reason, the following discussion of specification and prediction will apply to those kinds of items currently covered by MIL-F-8785B along with precision piloted tasks. In this sense, there are two objectives of design or procurement specification items:

- to guarantee required aircraft capabilities
- to guarantee pilot acceptance.

Ideally, if the basic postulate of flying qualities as stated in the Introduction has been followed, and an appropriate set of evaluation items has been established, these two specification goals are easy to achieve. All that is required for the first is to place inequality constraints on the evaluation measures, and for the second is to require favorable pilot ratings and comments for all test flight experience. There are two reasons why this does not currently suffice in practice:

- Appropriate sets of objective evaluation items have not been established and verified as sufficient for procurement.
- Specification items must be predictable during design and development; however, means of predicting performance and pilot acceptance for non-standard control and unconventional aircraft configurations are not fully developed.

For flying qualities as a subject to be fully responsive to the needs of current aircraft design and procurement agencies, these limitations on specification must be overcome. Two important achievements are required. First, in accordance with the above analysis of aircraft evaluation, it must be recognized that appropriate evaluation data sets should be established in terms of particular aircraft design and procurement objectives. As stated above, the guidance in doing this should be supplied by flight test and flight operational experience. Second, flying qualities as a subject must refrain from rephrasing all questions in terms of readily available or fashionable problem formulations, and instead respond by providing techniques for predicting specification compliance for all items deriving from the evaluation methods currently being used or being developed by flight test practice.

There are three kinds of evaluation data, objective, subjective, and analytical, that can be used to develop specification items to guarantee aircraft capability and to guarantee pilot acceptance. The ways in which these data are used, and the corresponding requirements of associated prediction methods will now be briefly summarized.

Specification and Prediction of Aircraft Capability

Open loop aircraft capabilities such as maximum attitude rates, rise times, mode phasing, trimmability and other such performance measures constitute much of the current MIL-F-8785B specification. Many of these items will remain appropriate for future aircraft procurement, and sufficient prediction methods exist based on transform, state variable, and time history representations of the aircraft and control system.

Since subjective evaluation is by definition a matter of pilot acceptance, the remaining approach to aircraft capability specification and prediction is by means of objective evaluation measures deriving from precision piloted tasks. For objective evaluation items developed for critical or representative flight tasks, specification items can be expressed in terms of statistics obtained by measurements in the time domain; that is, by observing what the aircraft is doing. Quantities such as mean and standard deviations of tracking errors, percent of time within allowable tolerances, and probabilities of exceedances are all easily measurable, but must also be predictable for any specification item based on them to be useful in aircraft design. Such prediction methods are now largely available, and given task and evaluation models as defined above, pilot-aircraft models can now be established and exercised to generate the required objective data predictions.

Specification and Prediction of Pilot Acceptance

The objective of MIL-F-8785B is to assure flying qualities that are "clearly adequate for the mission flight phase" when compliance is demonstrated. This is done by comparing one- or two-dimensional analytical or objective evaluation measures against inequalities (one-dimensional data) or boundaries (two-dimensional data) that have been validated to correlate with goodness of flying qualities defined in terms of Level 1 (clearly adequate), Level 2 (adequate but with increased pilot workload or degradation in performance), and Level 3 (safe flight, but inadequate flying qualities). It should be noted that the analytical and objective data correlated with Levels in MIL-F-8785B are performance measures of the open loop airplane only, consisting

of such items as airframe or augmented airframe frequency versus damping, roll-sideslip phase and amplitude, and controller gradient forces. All required evaluation data for MIL-F-8785B comparison are easily predicted.

This MIL-F-8785B approach to flying qualities specification by correlating one- or two-dimensional evaluation open loop data with Levels of flying qualities has been highly successful, and it is natural to extend this method to correlate closed loop objective evaluation data as well. The use of closed loop pilot-aircraft prediction models allows a much closer correspondence between specification items and design and procurement objectives. Supporting predictive means exist for a wide variety of general and representative evaluation data items that may be correlated in a one- or two-dimensional way with Levels of flying qualities.

This task and evaluation generality manifests itself in two fundamental ways:

- Transient or steady-state precision piloted tasks.
- Single or multiple task pilot activity for a given flight phase.

A survey of the four corresponding basic closed loop pilot-aircraft prediction models is presented in Reference 9.

So far, the discussion of pilot acceptance has concerned the use and prediction of analytical and objective evaluation data. In addition to these correlations, the direct use of pilot ratings should be considered. Experience has shown that current methods of training test pilots and introducing them to prototype or development aircraft leads to accurate predictions of acceptance by pilots of the resulting operational aircraft. For this reason, the final judgment of pilot acceptance of a given airplane rests with pilot ratings obtained during flight test programs. Specification of pilot acceptance simply becomes a matter of requiring acceptable pilot ratings during all flight test evaluation studies. During design, before the aircraft is available to a test pilot, correlations of evaluation measures can be used to predict acceptance as discussed above. There is, however, another approach: pilot rating prediction.

Several methods for predicting pilot ratings have demonstrated the ability to "predict" ratings for previously existing sources of experimental data. These methods postulate that the performance and workload aspects of the Cooper Harper rating scale are weighted by the pilot according to a linear functional, or can be predicted using optimal control and a pilot model performance index. A survey of the demonstrations of pilot rating prediction reveals that a number of underlying assumptions will

have to be justified before this attractive approach can gain acceptance as a prediction method, let alone as a basis of flying qualities specification.

First of all, pilot compensation models have been developed and validated by linear identification methods, and assume that the pilot operates as a time invariant continuous controller who generates control commands as a fixed blend of tracking error and its derivatives. The use of pilot models for rating prediction must assume that the pilot model parameters are related to workload, since aircraft tracking performance cannot be correlated with pilot model gains, leads, and lags. It must be further assumed that all workload aspects of flying qualities are manifested in the model coefficients.

Now consider pilot rating assumptions. It is assumed that aircraft performance can be normalized or calibrated in a manner that reflects a pilot's concern with adequate flying qualities. It is further assumed that the workload measure is a linear functional of the pilot model coefficients adjusted for optimal predicted pilot ratings, that the pilot rating is a linear functional of both performance and workload, and that the weighting coefficients are constants.

It should be noted that justification of these or similar assumptions is not required to extend pilot rating correlation (as contrasted to prediction) methods as practiced by MIL-F-8785B to include closed loop performance measures. Furthermore, it has been shown that by suitable choices of performance measures, tradeoff aspects of piloted experience can be identified, and correlations with pilot comments as well as pilot ratings can be obtained, Reference 8.

In summary, specification items must be developed in terms of the most meaningful evaluation items that flight test, flight simulation, and operational experience can evolve. Once these items are identified, at least the following conditions must be met for the item to be accepted as a procurement specification criterion:

- The specification item must be numerical.
- The specification item must correlate with pilot comments and pilot ratings.
- The specification item must be easily measured in flight test or flight simulation.
- The specification item must be reliably predictable by analytical means for use in early design and development.

- The method that predicts the specification item must be applicable in a completely standardized form that incorporates the most general models of the candidate aircraft required.

- The specification item must be valid for all current and acceptable aircraft, and must exclude poor or unacceptable aircraft.

FLYING QUALITIES PREDICTION METHODS

The preceding analysis of aircraft evaluation and specification has promoted the following two principles as a basis for developing the subject of flying qualities in a way responsive to the needs of advanced aircraft development:

- Evaluation and specification items should be developed for individual airplanes in a way that most generally reflects the operational requirements; flight test practice is a good guide for this.

- Specification items, to be useful in design, development, and aircraft improvement, must be supported by analytical prediction methods responsive to the full generality of the flight test evaluation items; it no longer suffices to limit flying qualities prediction to those quantities that can be calculated by steady-state, linearized, and transform methods.

Most pilot-aircraft prediction analysis has been concerned with a single axis tracking task and although this does not represent many flight phases well, it nevertheless is a useful description of a pilot's activity during precision maneuvers such as weapon delivery and landing. By extending flying qualities analysis from the consideration of open loop eigenvalues to the dynamically more complete model of the pilot's loop closure, approximations to the closed loop control can be obtained. From this standpoint, pilot — aircraft modeling work has been concerned with matching a simulator pilot's gain and phase as identified by linear means, and assessing the characteristics of the loop closure by methods of classical or optimal control theory.

The models developed by these means can also be used to predict the closed loop tracking statistics of a piloted task such as attitude stabilization in turbulence or following a randomly generated command. For such problems, the models show two important characteristics:

- Fixed form gain — lead — delay models agree within a few percent with flight simulation tracking error statistics.

- Motivated skilled pilots asymptotically trained achieve nearly identical tracking error statistics.

In this way, pilot — aircraft models are accurate predictors of what operational piloted aircraft can do during precise tracking. However, there are limitations on this approach, both in terms of the applicability of the model, and in terms of how the results have been interpreted. For example:

- Random tracking commands are difficult to relate to operational experience.
- Much pilot activity and concern is with the ability to make precision discrete changes in aircraft attitude, flight path or flight condition. These cannot be represented in a time-invariant manner.
- Flight simulation and pilot — vehicle analysis of a continuous tracking task are not representative of the full scope of aircraft flying qualities and in no sense "do the whole job."
- The common practice of obtaining pilot ratings during highly restricted precision tracking flight simulations has led to the incorrect assumption that these ratings reflect an overall rating of the aircraft dynamics simulated. Ratings obtained in this way reflect only compensation aspects of workload, and pilot estimates of performance are based only on experience during the flight simulation, which is not comparable to actual flight experience. For these reasons, pilot rating prediction methods that are adjusted to such simulation data cannot be regarded as validated predictors of overall ratings obtained from flight test.
- The reliance on single task continuous prediction methods assumes that flying qualities of an airplane can be fully studied by looking at each piloted task component separately. It is now widely recognized that the pilot has available limited attention, sensory, and motor information channel capacities which produce task interference effects that are severe limitations on performance in multi-task flight such as landing and weapon delivery.
- A practical limitation on the use and acceptance of single task time-invariant pilot — aircraft prediction methods has been the failure of most studies on the subject to subordinate the specific model components to the overall concerns of what the aircraft does, and how well a pilot can make it perform.
- Reports on pilot-aircraft methods often present elaborate arguments concerning model parameters rather than derive time domain statistics and properties that can be related to flight simulation and flight test. As long as this tendency persists, the aircraft control design community will continue to regard pilot — aircraft methods as simply "pilot modeling," an esoteric subject not fully responsive to design and development requirements.

It was indicated above that a pilot's control of an airplane can be conveniently classified in generic terms of whether his particular flight task is continuous steady-state or transient and intermittent, and whether he is faced with only one attention demand, or if several independent activities are under his control. The authors have recently completed a study for the USAF Flight Dynamics Laboratory that demonstrates

how these categories of flying qualities problems can be studied in accordance with the above principles. Reference 3 presents a comprehensive account of these methods, and Reference 9 gives a brief summary of these results along with examples of how flying qualities prediction methods can be developed.

REFERENCES

1. Twisdale, Thomas R., and Franklin, David L., Tracking Test Techniques for Handling Qualities Evaluation, AFFTC-TD-75-1, Air Force Flight Test Center, Edwards Air Force Base, California, May 1975.

2. Moorhouse, D.J., The History and Future of U.S. Military Flying Qualities Specifications, American Institute of Aeronautics and Astronautics Paper 79-0402, Seventeenth Aerospace Sciences Meeting, New Orleans, Louisiana, January 1979.

3. Onstott, E.D., and Faulkner, W.H., Prediction, Evaluation, and Specification of Closed Loop and Multiaxis Flying Qualities, AFFDL-TR-78-3, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio, February 1978.

4. Harper, R.P., and Cooper, G.E., A Revised Pilot Rating Scale for the Evaluation of Handling Qualities, Cornell Aeronautical Laboratory Report No. 153, September 1966.

5. Anonymous, Military Specification, Flying Qualities of Piloted Airplanes, MIL-F-8785B, August 1969.

6. Chalk, C.R., et al, Background Information and User Guide for MIL-F-8785B (ASG), "Military Specification — Flying Qualities of Piloted Airplanes," AFFDL-TR-69-72, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio, August 1969.

7. Neal, T.P., and Smith, R.E., An In-Flight Investigation to Develop Control System Design Criteria for Fighter Airplanes, Volumes I and II, AFFDL-TR-70-74, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base, Ohio, December 1970.

8. Onstott, E.D., and Faulkner, W.H., "Discrete Maneuver Pilot Models for Flying Qualities Evaluation," Journal of Guidance and Control, Vol. 1, No. 2, March-April 1978.

9. Onstott, E.D., and Faulkner, W.H., Definition and Interpretation of Flying Qualities, NOR-79-22, Northrop Corporation, Hawthorne, California, March 1979.

AAA TRACKING WITH DELAYED, DISCRETE-TIME FEEDBACK

by
Arye R. Ephrath
Department of Electrical Engineering & Computer Science
University of Connecticut
Storrs, Connecticut 06268

Abstract

In most target-tracking experiments the operator controls a reticle sight to position the cross-hairs on the simulated target. Our study focused on the human's ability to control not the sight, but the projectile path--i.e., the gun itself.

Six trained subjects tracked a simulated target airplane on a CRT during a straight-and-level pass. The subjects tracked without a reticle sight but were provided with feedback of miss-distance in the form of simulated tracer rounds on the CRT screen. Three experimental conditions were investigated: (a) the full stream of tracers is visible; (b) only the"tip" of the stream (i.e., the one tracer-round closet to the target) is visible; (c) the full stream of tracers is visible, plus lead-angle information in the form of a reticle sight.

The experimental results are presented and comparisons are made among the tracking performance characteristics under the different conditions. Some conjectures and suggestions for further research are included.

Introduction

In most Anti-Aircraft tracking experiments, the operator controls a reticle sight to position a cross-hairs on a simulated target. The hypothetical "gun" is assumed to point ahead of the target, the lead angle computed in some manner such that a round fired now would subsequently impact the target. The projectile's ballistic path is never seen by the subject.

In the experimental investigation reported here, the subjects were provided with feedback of hit performance (miss-distance) in the form of simulated tracer rounds. From analytic and modelling viewpoints, tracers are quite interesting as the error information that they provide is discrete, time-delayed, and non-stationary (as the time-delay is range-dependent). In addition, the human operator must do his own lead-angle prediction, thereby making the tracking task quite difficult. To document manual tracking performance with tracers, and to provide the necessary data-base for future analytic work, we conducted a set of experiments in which the target (a low-flying aircraft) and the tracers were presented to the subjects as images on a CRT screen. The subjects' control inputs were sampled and stored for off-line analysis of tracking performance.

1. Experimental Set-Up

Our experimental facility consists of a DEC PDP-11/60 computer, coupled with a VS-60 High Performance Graphics System. The graphics unit provides refreshed vector images of extremely high quality on a 21-inch CRT screen. To sample the

subjects' control inputs we use Air Force-type control grips which are wired into the computer's processor via A/D convecters. The control signals are digitized with a 12-bit resolution (one part in 4096).

Each subject in our experiments was seated in front of the 21-inch CRT screen. With the right hand, the subject manipulated a force-actuated (isometric) control stick which generated a rate command to the simulated gun pointing angle. The control stick gain was set to the value which had been judged "most comfortable" during the shakedown runs.

The forcing function $\theta_T(t)$ was the position of a delta-shaped target image during a straight-and-level flyby. The target was assumed to be at a constant altitude of 200 ft., moving from left to right at a constant velocity of 200 mph, with a cross-over range of 700 ft. At a display-gain value of 1.0 these parameters provided the best trade-off among image size, $\dot{\theta}_T$ and $\ddot{\theta}_T$ (target angular velocity and angular acceleration, respectively).

The subject was assumed to sit between two AA guns, 4 ft. apart. These simulated guns produced two parallel tracer streams on the CRT screen. Tracers were produced at a constant rate of 5 pairs/second throughout the run; the subjects had no control over the simulated firing rate. The muzzle velocity of the tracers was set to 2000 fps and they moved in a ballistic trajectory under the influence of earth gravity. Secondary effects due to air resistance, gyroscopic precession, spin and wind were neglected. The resulting equations of motion of tracer i are, then:

$$x^i(t) = x^i(0) + V_p \cdot \cos\varepsilon^i \cdot \sin\gamma^i \cdot t$$
$$y^i(t) = y^i(0) + V_p \cdot \sin\varepsilon^i \cdot t - 1/2\, gt^2$$
$$z^i(t) = z^i(0) + V_p \cdot \cos\varepsilon^i \cdot \cos\gamma^i \cdot t$$

where (x,y,z) is a cartesian coordinate system centered at the subject with x parallel to the flight-path of the target (positive to the right), y vertical (positive up) and z completing a left-handed triad; V_p is the muzzle speed of the tracers (2000 fps), ε^i and γ^i are the values of the gun's elevation and azimuth angles, respectively, when tracer i was fired, t is tracer i's time of flight and g is the gravitional acceleration (32.2 fps/sec).

Because of the inherent limitations of a flat, two-dimensional screen in providing depth cues, tracers were "terminated" (i.e., their images were removed from the screen) when they reached a slant range equal to the target's slant range. Thus, to minimize tracking error,† or "hit score", the subjects had to point the "gun" in such a way that the <u>end</u> of the tracer-stream coincided with the target's image.

The positions (in inertial space) of the simulated target and of each tracer were computed in real time. The corresponding coordinates of their images on the

† In the present context "tracking error" is defined as the angular distance between the target and tracer stream endpoint as seen by the subject. "Pointing error" refers to the angle between the gun's present position and the perfect position required for a "hit" after a time-of-flight. Thus, present pointing errors give rise to future tracking errors.

graphics screen were updated at a rate of 20/second. The subjects' control inputs in both azimuth (x) and elevation (y) axes were sampled at the same rate of 20/second and stored in individual disk files for off-line processing. Each run lasted about 25 seconds and produced 512 datum points of each sampled variable.

Three experimental conditions were incorporated in the design, as follows:

Condition A: Subject tracks the target with a stream of tracers, plus a set of cross-hairs which provides lead-angle information (Fig. 1).

Condition B: Subject tracks the target with a stream of tracers. No lead information is provided (Fig. 2).

Condition C: Subject sees only the end-point of the stream of tracers, displayed as a cross (Fig. 3).

Under Condition A, lead angle information was computed in real time based on the target's present position and velocity. This information was presented to the subject in the form of crosshairs on the screen: When the crosshairs were aligned with the target's image, the tracer round just then leaving the gun would hit the target in a few seconds (time-of-flight). The crosshairs were not stationary on the screen, since the lead angle is a function of the target's present range, relative velocity, etc. The subject thus receives dynamic information as to where to point the "gun".

Under condition B, there was no reticle, gun-sight or cross-hairs display. The subjects tracked the target in much the same way as a firefighter would control the stream of water coming out of his hose. Under Condition C, the stream of tracers was not visible on the screen; a visible cross denoted the instantaneous end-point of the "invisible" tracers-stream and the subject's task was to place this cross on the target image.

Each subject received sufficient training in tracking under the above conditions, one condition at a time. When a subject's performance under an experimental condition reached the desired level he was presented with 15 identical formal runs under that condition before proceeding to the training phase under the next condition. Each subject thus participated in a total of 45 formal runs. The order of presentation of the experimental conditions was counter-balanced among the subjects.

Six subjects, all University of Connecticut students, participated in the formal experiment. All were male, right-handed, with normal vision, and between the ages of 21-25. Five were Air Force ROTC cadets; two had flying experience, five had sharp-shooting experience.

All of the subjects were highly motivated. To further increase their motivation, however, an imaginary 5' x 9' "bulls-eye" was assumed at the center of the 30' x 15' target aircraft: Only tracers which terminated at this reduced bulls-eye were counted as "hits". At the end of each run the number of hits scored was displayed on the screen (no other performance feedback was provided) and the subject was encouraged to try to do better.

Five variables were sampled and stored during each formal run: The (x,y,z) coordinates of each tracer as it reached a slant range equal to the target's, and the control stick inputs in azimuth and elevation. Note, however, that the variable of greatest interest, namely, the tracking error or miss distance, was not recorded. This variable could be reconstructed from the recorded data off-line and thus, in an

attempt to minimize disk storage requirements, was not stored explicitly.

Tracking error (or, more accurately, "miss angle") could be computed from the coordinates of the "last" tracer (that is, the tracer which had just reached the target's slant range): As the target's trajectory was deterministic its coordinates could be computed and compared with the coordinates of the corresponding "Last" tracer, to yield angular error. This approach, however, was abandoned for three main reasons:

1. There is no guarantee that a tracer would reach a slant range equal to the target's every sampling interval (1/20 sec.).

2. At a mussle velocity of 2000 fps a tracer travels about 100 ft. during each sampling interval; the target travels approximately 15 ft. in the same time. Thus a tracer's coordinates are accurate to within $\pm$ 50 ft., and the target's to within $\pm$ 7.5 ft., yielding inaccuracies and noise in the resulting miss angle. Although these inaccuracies may be minimized by some curve-fitting interpolation technique, the procedure is complicated and cumbersome.

3. This technique yields a value for the miss angle, $e(t)$, at time t when a tracer reaches the target's slant range. As far as the subject's performance is concerned, however, the error is attributable to the time $(t-\tau)$ when that tracer left the gun, τ being the tracer's time of flight. The time-shift τ is not constant, and it is a function of the target's slant range. As a consequence, translating the error sequence $e(t)$ in time to generate the sequence $e(t-\tau)$ may prove to be a complicated and cumbersome proposition.

Instead, a different algorithm was employed to compute the gun-pointing error: Based on the target's trajectory (which was deterministic) a sequence of perfect lead-angles was generated off-line, i.e., angles at which the gun should be pointing in order for the tracer just-then being fired to hit the target. These values were subtracted from the corresponding (recorded) actual control inputs to yield the desired pointing-error sequence.

When the target's motion is completely known (as it was in our experiments), and when the computation of the lead-angles sequence is based on that known target motion (as it was in this case), then it can be shown that the gun-pointing-error sequence is equivalent to the sequence $e(t-\tau)$ of item (3) above. In other words, a gun-pointing-error e_i at the time a tracer leaves the gun gives rise to an equal miss-angle when that tracer reaches the target's slant range, τ seconds later.

As a check, the two methods (that of "miss angles" and that of "pointing errors") were compared. They proved to be equivalent, apart from the time-varying time-shift τ. However, the lead-angle, gun-pointing-error method produced noticeable smoother results. All the results shown in this paper in graph form are gun-pointing-errors, derived by the lead-angle method.

2. Results and Discussion

As a first step in analyzing the results the data were aggregated, i.e., point-by-point ensemble means and mean-squares were computed. We utilized a sequential approach to the computation, by which the mean m^j_{k+1} of the j-th point $(0<j\leq 512)$ of k+1 runs was computed from m^j_k by

$$m^j_{k+1} = (m^j_k \cdot k + z^j_{k+1}) \; / \; (k+1)$$

where z^j_{k+1} is the value of the j-th point in the (k+1)-th run, and $m^i_1 = z^i_1$, i = 1, 2,...,512.

The mean-square q^j_{k+1} was computed similarly from q^j_k by

$$q^j_{k+1} = (q^j_k \cdot k + (z^j_{k-1})^2) \; / \; (k+1)$$

and the unbiased estimate of the variance, v^j_k, was then found from

$$v^j_k = (q^j_k - (m^j_k)^2) \cdot k/(k-1)$$

Outliers were identified and removed prior to the averaging process: As the CRT screen subtended a visual angle of approximately $\pm$ 13°, tracking errors of magnitude larger than 13° indicated loss of the target's image. We set the cut-off point, somewhat arbitrarily, at 3 seconds: Any run containing a tracking error (in either axis) in excess of 13° for more than 3 seconds was labled an outlyer and removed from the statistics. A total of five runs were thus removed (1 from Condition "B" and 2 each from Conditions "A" and "C"). The surviving runs are shown in Figs. 4 - 6. As expected, tracking performance was best under Condition "A" (presence of lead-angle information). Unexpected, however, were some characteristics of the data, such as:

1. The tracking performance with only end-point information is much worse than the performance when the full tracer-stream is visible: Peak azimuth errors under Condition C are three times larger than the corresponding error under Condition B. This fact came to us as an unexpected surprise, as the tracer-stream had not been thought to contain useful information apart from its end point. Obviously, it does, and this may be a useful area for future investigation.

2. The runs commenced with a non-zero initial error, both in azimuth and in elevation, of about 1°. Invariably, this initial error effected an overshoot of about 2° under conditions B and C (there was no overshoot with lead-angle information), indicating perhaps some lack of tightness in the control loop even in this region of so-called "good tracking".

3. The error variance exhibits an interesting behavior: Under Condition A (lead-angle information) it is low throughout, except for some increase in the cross-over region. Under conditions "B" and (especially) "C", however, the variance shows a tendency to increase with time (Figs. 5-6). This tendency was especially evident in individual subjects' statistics where this trend was not masked by inter-subject variability (Fig. 7).

The phenomenon of an increased variance without a correlated change in the mean is typical of unstable oscillatory behavior, such as oscillations with increasing amplitude. Keeping in mind the nature of the task, namely, tracking a moving target with a stream that is in effect a long and flexible "hose", one can suspect such instability as the "hose length" increases. It may be rooted in the increasing time-delay, which effects a limit-cycle behavior equivalent to pilot-induced oscillations (PIO). It is less clear why this behavior is more pronounced under Condition "C", with only end-point information, as compared to Condition "B" (full stream). The question of the particular mechanism of the damping, evident when the full tracer stream is visible, as well as the causal relationship between the magnitude of the time-delay and the instability, remain yet to be explored.

6. Acknowledgements

The software for the simulation used in this research was designed and written by Mr. Donald Geller, who did an excellent job. The experiments were competently managed and administered by Ms. Barbara Chernoff. This project was funded under AFOSR Grant 77-3126.

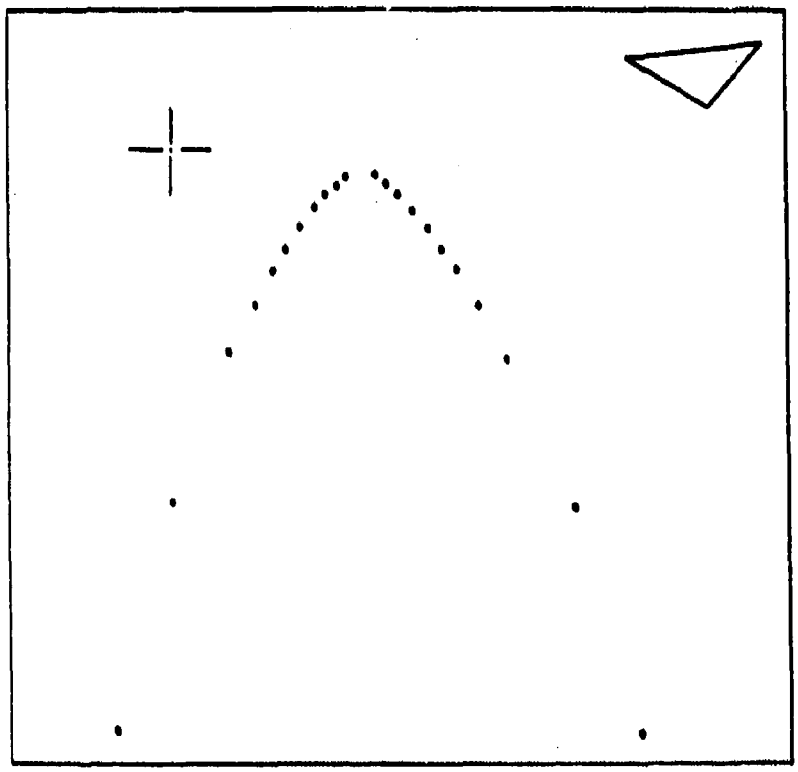


Fig. 1. Display Format for Experimental Condition "A" (Tracer Stream and Lead-Angle Information).

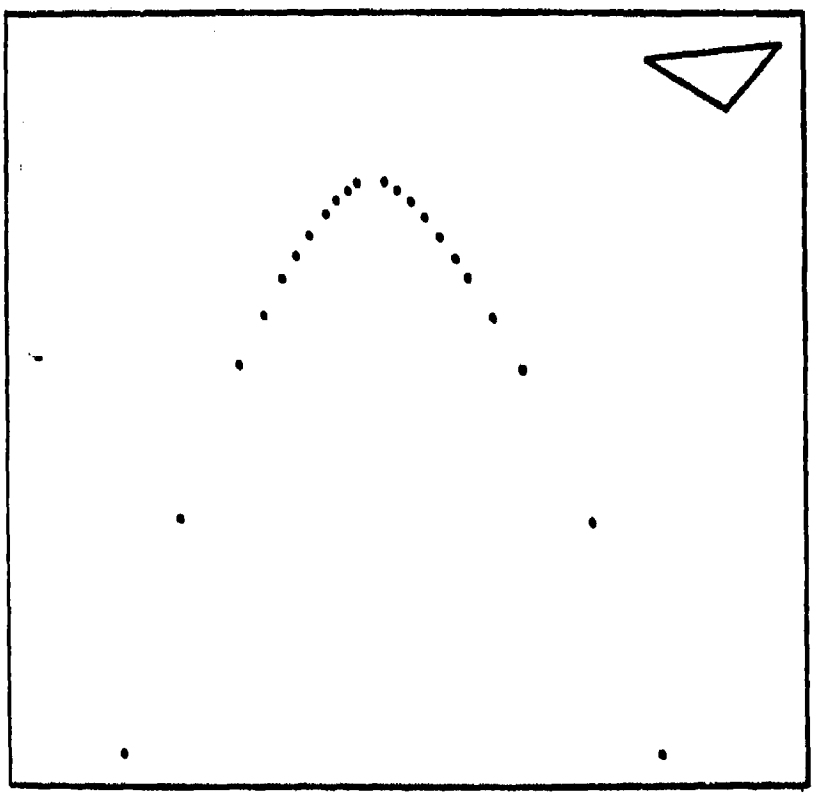


Fig. 2. Display Format for Experimental Condition "B" (Tracer Stream, No Lead-Angle Information).

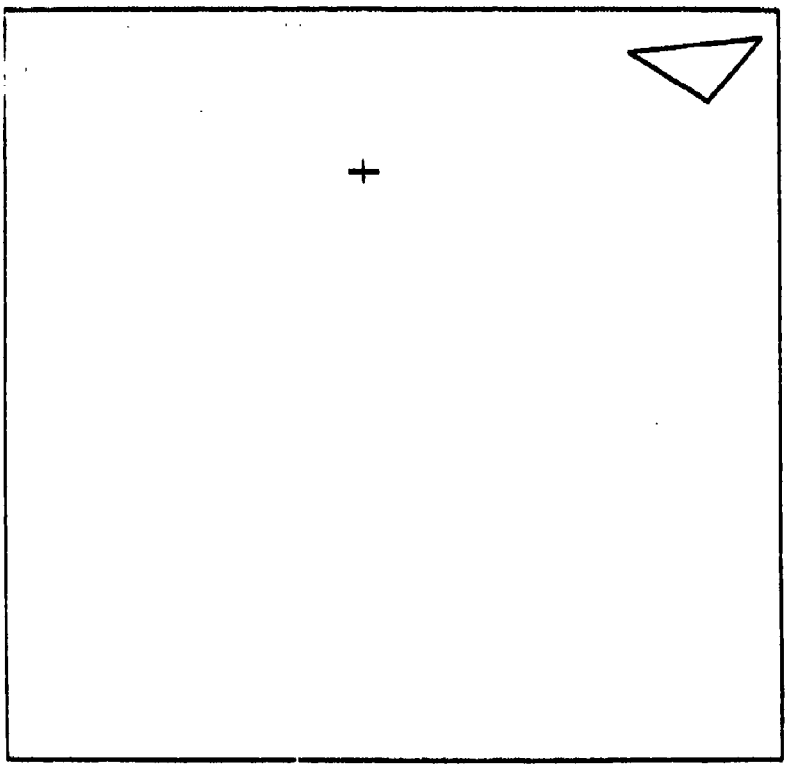
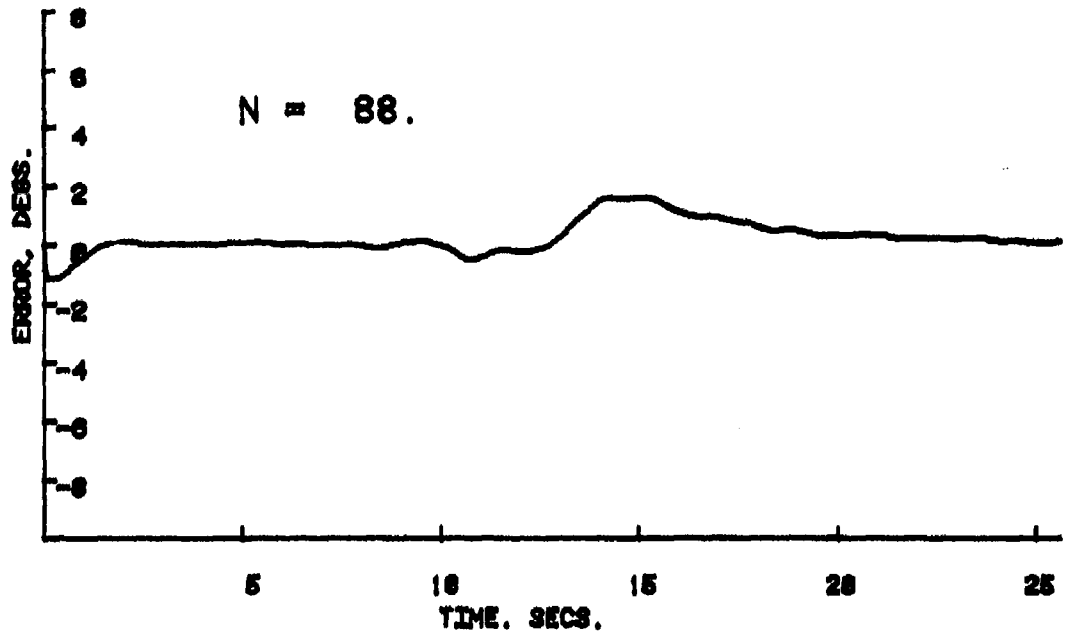


Fig. 3. Display Format for Experimental Condition "C" ("Invisible" Tracer Stream, Visible End-Point).



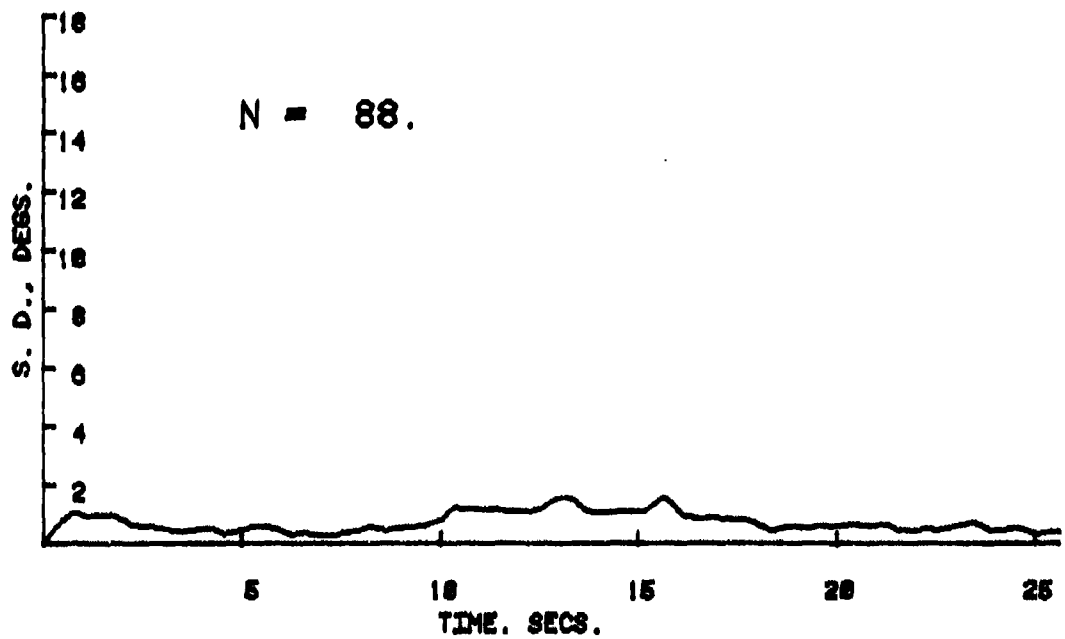
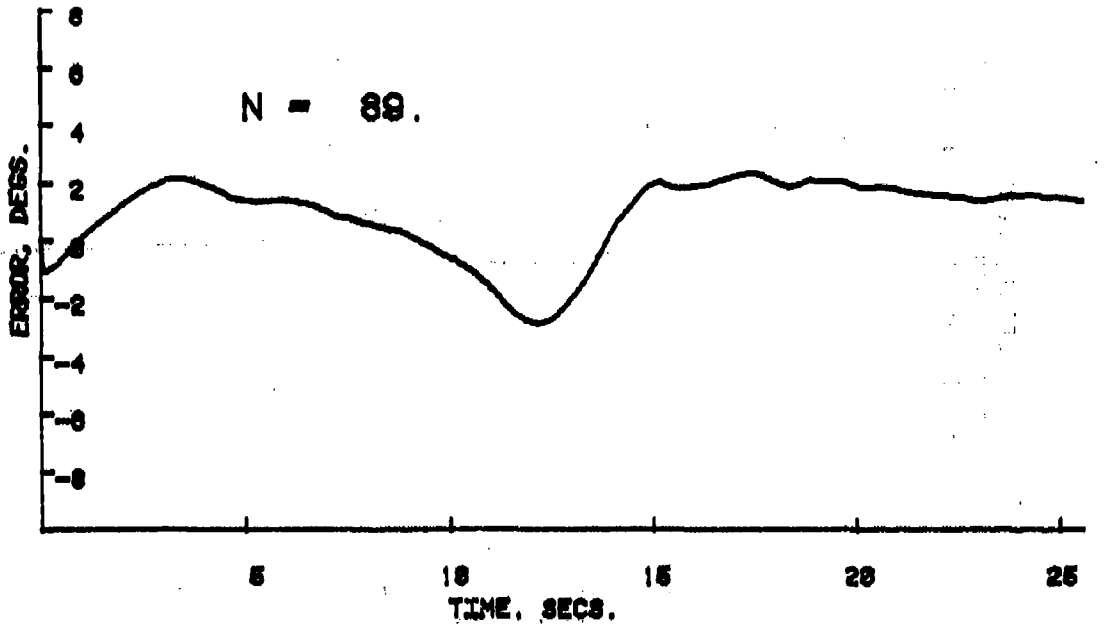


Fig. 4. Ensemble Statistics of Azimuth Error, Condition "A".



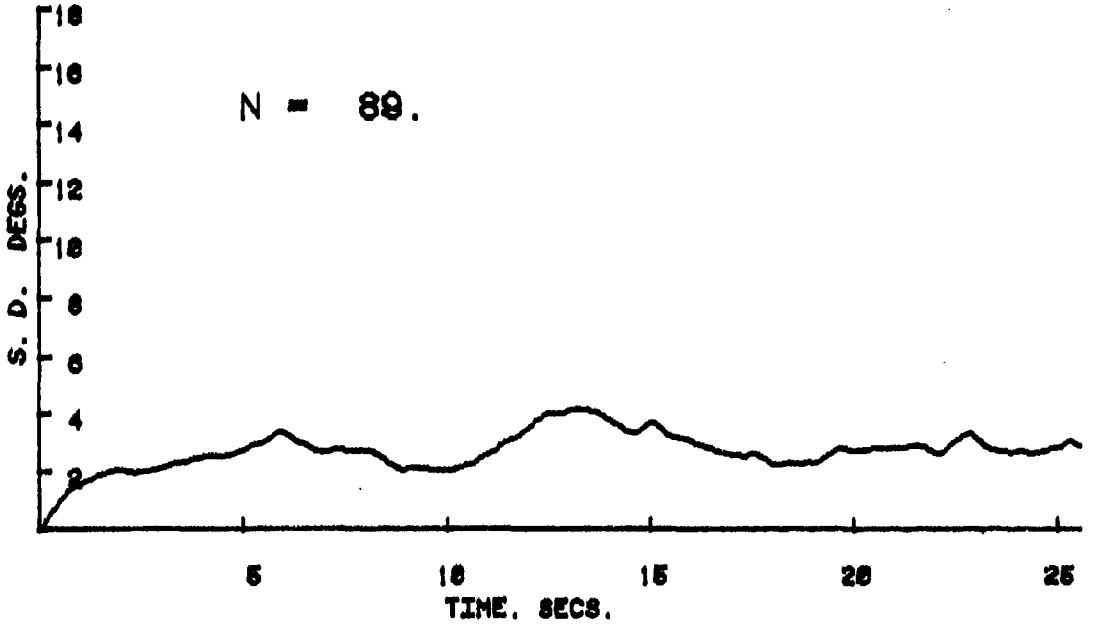
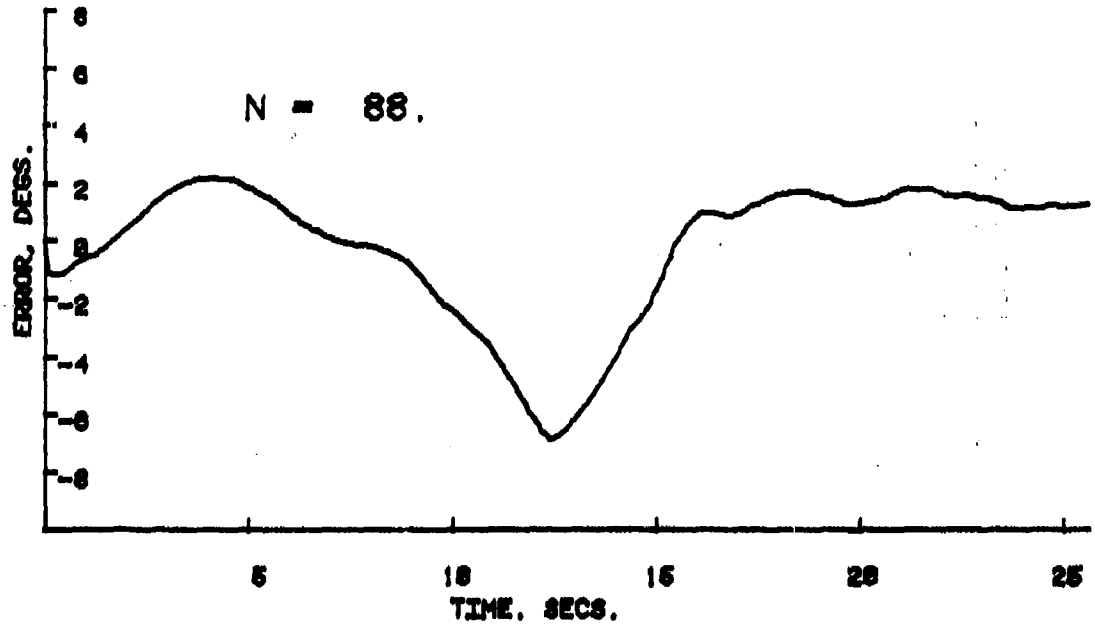


Fig. 5. Ensemble Statistics of Azimuth Error, Condition "B".



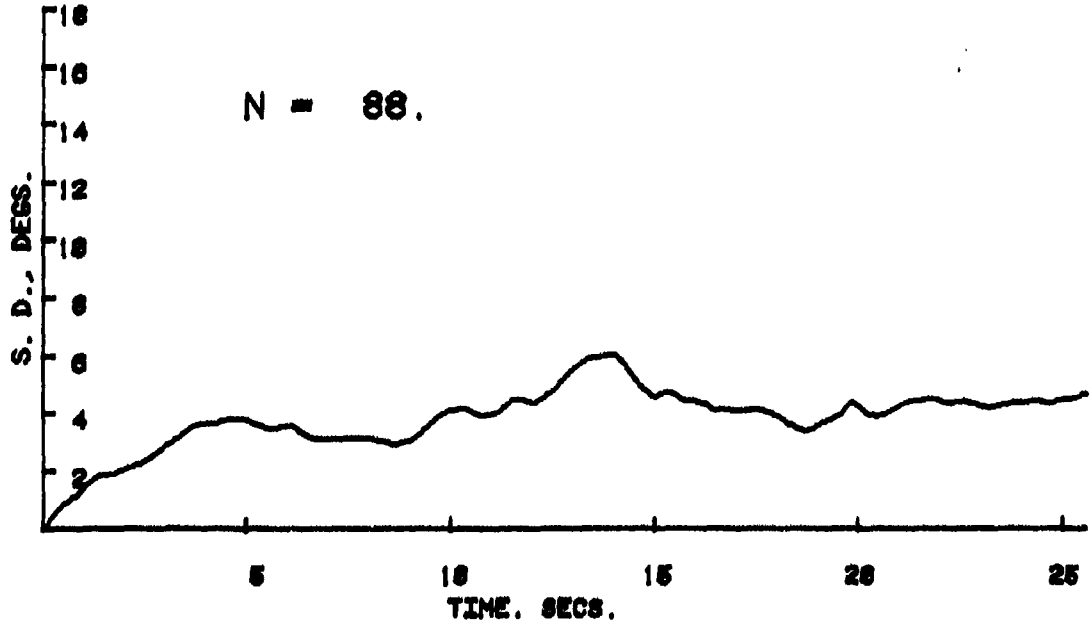
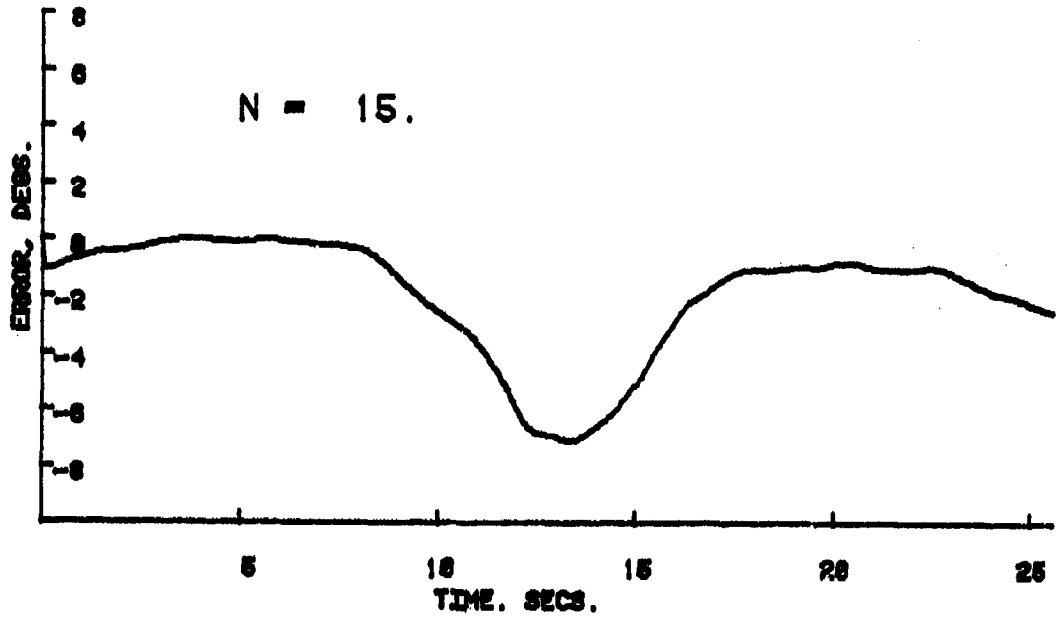


Fig. 6. Ensemble Statistics of Azimuth Error, Condition "C".



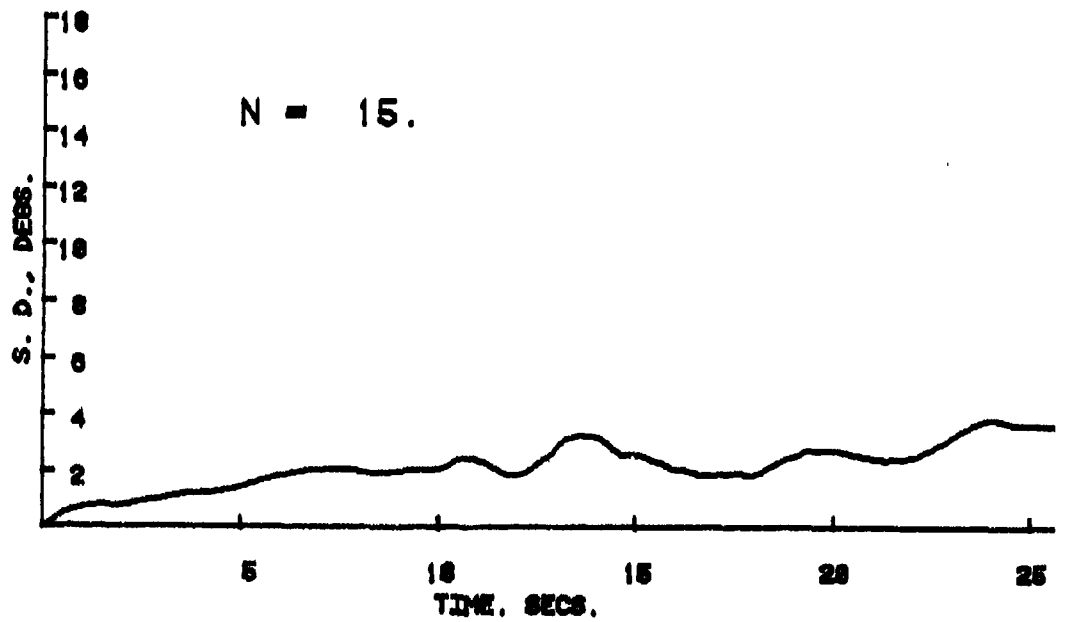


Fig. 7. Ensemble Statistics of Azimuth Error, Subject DG, Condition "C".

ANALYSIS OF A TRACER FIRE SYSTEM AS A MAN-MACHINE PROBLEM

D. W. REPPERGER M. M. VIKMANIS

Aerospace Medical Research Laboratory
Wright-Patterson Air Force Base, Ohio 45433

ABSTRACT

A study is conducted on data from a tracer fire system (Anti-Aircraft Artillery Simulation). Treating this problem within the context of a man-machine system, the effective closed loop capability is studied by determining critical characteristics of the target forcing function which specifies when the human has control of the total man-machine system. A decision rule based on properties of the closed loop tracking error is determined and used to study the limitations of the man-machine interaction with respect to different target forcing functions. A comparison is made using this approach between a good and poor tracker. The effect of tracking with and without a lead angle computer is studied for one of the three trackers considered here.

Introduction

In the study of a manned weapon within the context of a man-machine system, it is desired in some applications to obtain a physical description of the effective capability of the closed loop interaction. The approach considered here will involve the determination of a critical boundary to ascertain upper and lower bounds of special characteristics of the target forcing function which may lead to different types of tracking behavior. Such a method has been discussed in [1] and allows a procedure to determine limits of tracking capability in a manner analogous to a critical task situation [2]. It is the purpose of this study to apply this heuristic method as described in [1] to a tracking task of some complexity to develop results which may have ease of interpretation in terms of real world applications.

The Tracer Fire Problem

The man-machine problem considered here is termed a tracer fire problem. Figure (1) illustrates an analogy to this different type of tracking task and illustrates how the human interacts in this tracking situation with various types of information sources. In Figure (1), the analogous task considered here involves a human given the problem of trying to maneuver a hose which contains a stream of water. The subject manipulates

the hose in such a way as to direct the stream of water to strike a target at a time h seconds from when the water leaves the nozzle of the hose. The target lies in a direction perpendicular to the direction of the stream of water. For a direct hit, the stream of water arrives at the end point in h seconds; therefore, the goal of this tracking task is for the target also to arrive at this termination point in the same amount of time. The strategy for the subject, therefore, is to aim the hose in front of the target which is moving at right angles to its path. The subject can obtain rate information from the stream of water as to how the error in Figure (1) is nulled as a result of movements of the hose nozzle. As is seen in this diagram, three types of information are available to the tracker. The subject has the position of the nozzle as one source of information, the visual display of the stream of water is a second source of information, and the error perceived (e(t)) at the intersection of the target and the stream of water is a third source of information.

Figures (2a-c) represent elements of a tracer fire system which is, in many ways, analogous to the tracking problem depicted in Figure (1). Figure (2a) crudely represents the tracer fire system in block diagram form. The human operator is involved in controlling the simulation of a gun system on the ground. He has available display information of various kinds (Figure (2c)). The human generates a crank response which changes the position (Θ_B) of the gun which tries to follow the target flyby Θ_T. From the display, the human obtains diverse information such as gun position, tracer position as a function of time, and the position of the target aircraft. Figure (2c) illustrates this display. The tracking problem of interest to the human operator is to null the error e_1-e_2 which produces a direct hit on the target at the target range.

Figure (2b) illustrates a more detailed block diagram representation of Figure (2a). In Figure (2b), the human operator has available the error e_1 between the gun barrel and the plane. The tracker also has available the error e_2 which is the difference between the gun position and point at which the tracers intersect the target slant range. In Figure (2b), the gun system with a rate control system is modeled as a c/s system. The right-hand optics display that the human observes has Θ_T (the aircraft target), Θ_p (the display of the tracers), and Θ_B (the position of the gun barrel) available to the human. It is from these various sources of information that the human will choose a strategy to track this difficult type of problem.

For the experiment considered here, three trackers were used in this analysis. Two modes of operation were also considered (lead angle computer and manual tracking). In the manual tracking mode, the human viewed the target on the display and tracked the optical display errors as indicated in Figure (2c). In the lead angle computer mode, the lead angle computer generates a signal which moves the gun in a direction

proportional to the target plane movement across the display. That is, the computer generates a lead angle based on a constant target velocity and range during the projectile time-of-flight. The tracking task in this mode of operation is somewhat easier because the human only has to null out the error between the center line of the sight and the target; this is a standard compensatory tracking task. A procedure for obtaining an analytical technique to study this type of tracking problem is presented next.

An Analytic Procedure for Studying this Tracking Task

A heuristic approach for analyzing this type of tracking task can be obtained by studying Figures (3a-c) and Figure (4). Figure (3a) represents typical trajectories which represent the Azimuth component of the target trajectory. Figures (3b-c) represent, respectively, the velocity and acceleration of the target trajectory. In Figure (4), the velocity and acceleration of f(t) are plotted parametrically with time. A circle of radius R_1 is drawn in Figure (4) to delineate two regions (R_1 and R_2). It is assumed here (as in [1]) that region 1 which is inside of the circle is the relatively easy segment of the tracking task since the target flyby is slow in terms of velocity and acceleration. The region R_2 (which is outside the circle) represents the time period when the forcing function is difficult to track. By parametrically studying the forcing function in Figure (4), the regions R_1 and R_2 will describe when the target forcing function becomes too difficult to track.

When observing Figures (3a-c), Figure (5) is then constructed to be a qualitative representation of how a human behaves during this type of tracking task. The first and third segments of the tracking task in Figure (5) represent the regions where tracking is easy. This is seen to be true from Figures (3b-c) when the velocity and acceleration of f(t) are relatively small. During the middle third of the tracking task, however, the velocity and acceleration of the forcing function are large, which is representative of a fast moving target. It is during this middle segment when it is difficult to track and the human will tend to lower his gain; a qualitative description of this behavior may be termed "regression period." This is analogous to crossover regression reported in [3].

The manner in which it is easy to detect when a human moves from region 1 to region 2 can be observed in the error phase plane (Figure 6). During the easy portions of the tracking task, the error signal oscillates about the origin in a small ellipse. When the closed loop error diverges significantly from the origin, the human has regressed and here it is assumed that he has entered a different mode of tracking behavior. Figure (7) illustrates the two regions considered. It is necessary in this analysis to determine the boundary of the ellipse. Several methods exist in the determination of this boundary. Figures (8a-c) illustrate a heuristic method considered in this study.

In Figure (8a-c) the miss distance of error was plotted versus time. By definition (Figure (9)):

$$e_{\text{Miss Distance}} = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2 + (z_1-z_2)^2} \quad (1)$$

$$= R\sqrt{d_1^2 + d_2^2 + d_3^2} \quad (2)$$

where

$$d_1 = \text{Sin}\ \phi_1\ \text{Cos}\ \theta_1 - \text{Sin}\ \phi_2\ \text{Cos}\theta_2 \quad (3)$$

$$d_2 = \text{Sin}\ \phi_1\ \text{Sin}\ \theta_1 - \text{Sin}\ \phi_2\ \text{Sin}\theta_2 \quad (4)$$

$$d_3 = \text{Cos}\ \phi_1 - \text{Cos}\ \phi_2 \quad (5)$$

Where R is slant range, ϕ_i and θ_i $(i=1,2)$ are the spherical coordinates, and e the difference in position between the point at which the tracers disappear and the aircraft, which represents total error distance. From Figure (8a), a nominal value of error Miss Distance is determined when the man-machine system is tightly closed. When the closed loop error exceeds three times the nominal amount, the time t_1 is determined. Projecting this time down on the plots of $\dot{f}$ and $\ddot{f}$ yields values of $\dot{f}(t_1)$ and $\ddot{f}(t_1)$ when the tracking task is very difficult. In this manner the regions of the $\dot{f}$-$\ddot{f}$ plot of Figure (4) can be approximated by plotting the critical region from a catalog of values obtained from this procedure. The results of this analysis procedure are presented next.

Results of This Analysis Procedure

Using this approach on five different flyby functions (termed 2x2 flyby, zig-zag, jink, weapon delivery, and recon), the results are displayed in Table I. Three trackers were chosen (#33, #80, and #93) and they were classified as best tracker, mediocre tracker, and worst tracker based on performance analysis using measures of performance such as time on target. In Table I, these results are displayed for the lead angle computer tracking mode.

Table II considered all five flyby functions in both the Azimuth and Elevation axis for the same three trackers without the lead angle computer. It is noted that in Table II, the order of performance hierachy of the trackers differed with the lead angle computer as compared to no lead angle computer.

Figures (10a-b) illustrates the plots obtained from this procedure. In plot (10a), a comparison is made between a good and poor tracker. Since five different flybys were considered, the region is enclosed within five sample points to determine the critical boundary. It is noted that the good tracker has a larger region of possible forcing functions he can track as compared to the poor tracker, i.e., the class of possible target trajectories is quantified in Figure (10a) and each tracker is explicitly quantified by the trajectories he can track. Also, a technology of possible target trajectories is displayed which is beyond the capability of the manned gun system. Also obtained from this analysis procedured is a quantitative definition of the class of possible trajectories the manned gun system can follow.

Figure (10b) illustrates a comparison of a lead angle computer tracking mode versus a non-lead angle computer tracking mode for tracker #93. One can see from this figure the advantage of using the lead angle computer and how it improves the capability of the man-machine system. One notes that the lead angle computer extends the tracking region substantially by changing the type of tracking task. Also from this figure one notices that the class of flybys where the lead angle computer improves the capable tracking region and quantitatively defines when the lead angle computer is advantageous over optical (manual) tracking. Also from a tactics point of view, Figures (10a-b) illustrate what type of tactics are necessary to beat the man-machine system. The class of tactics which beats the manual system appears less than for the lead angle computer. We therefore have a quantification of these regions.

One note of caution is mentioned now. From Figures (10a-b) we have a crude approximation to Figure (4). In order to determine Figure (4) more accurately, it would be necessary to run a large number of different target forcing functions to sweep out the phase plane. In addition, the boundary of Figure (4) is a mean of replications considered in the analysis. Actually there exists a 1σ boundary about this region and Figures (10a-b) are a small sample estimate of the regions of Figure (4). The critical boundary is actually a random variable and an envelope exists around this boundary. Therefore, in the validation of and use of this method for prediction, one must consider the fact that the critical boundary is a mean value and a distribution about this mean value is necessary to characterize its density (see [1] for more specific details).

Summary and Conclusions

A decision boundary is obtained which describes the regions of the target forcing function which the human can track or and regions where tracking is difficult. A critical boundary is constructed based on five target trajectories and with three trackers, and with two modes of tracking (lead angle computer and no lead angle computer). This procedure allows a tactics technology to be developed concerning the capability of the man-machine system.

References

[1] Repperger, D.W., S. Ward, E.J. Hartzell, B. Glass, and W.C. Summers, "An Algorithm To Ascertain Critical Regions of Human Tracking Ability," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-[illegible], No. 4, April 1979, pp. 183-196.

[2] Jex, H.R., J.E. McDonnell, and A.V. Phatak, "A Critical Tracking Task For Man-Machine Related to the Operators Effective Delay Time I: Theory and Experiments With a First Order Divergent Controlled Element," NASA-CR-616, October 1966.

[3] McRuer, D., D. Graham, E. Krendel, and W. Reisner, "Human Pilot Dynamics in Compensatory Systems-Theory, Models, and Experiments with Controlled Element and Forcing Function Variations," AFFDL-TR-65-15, July, 1965.

AN UNUSUAL TRACKING PROBLEM

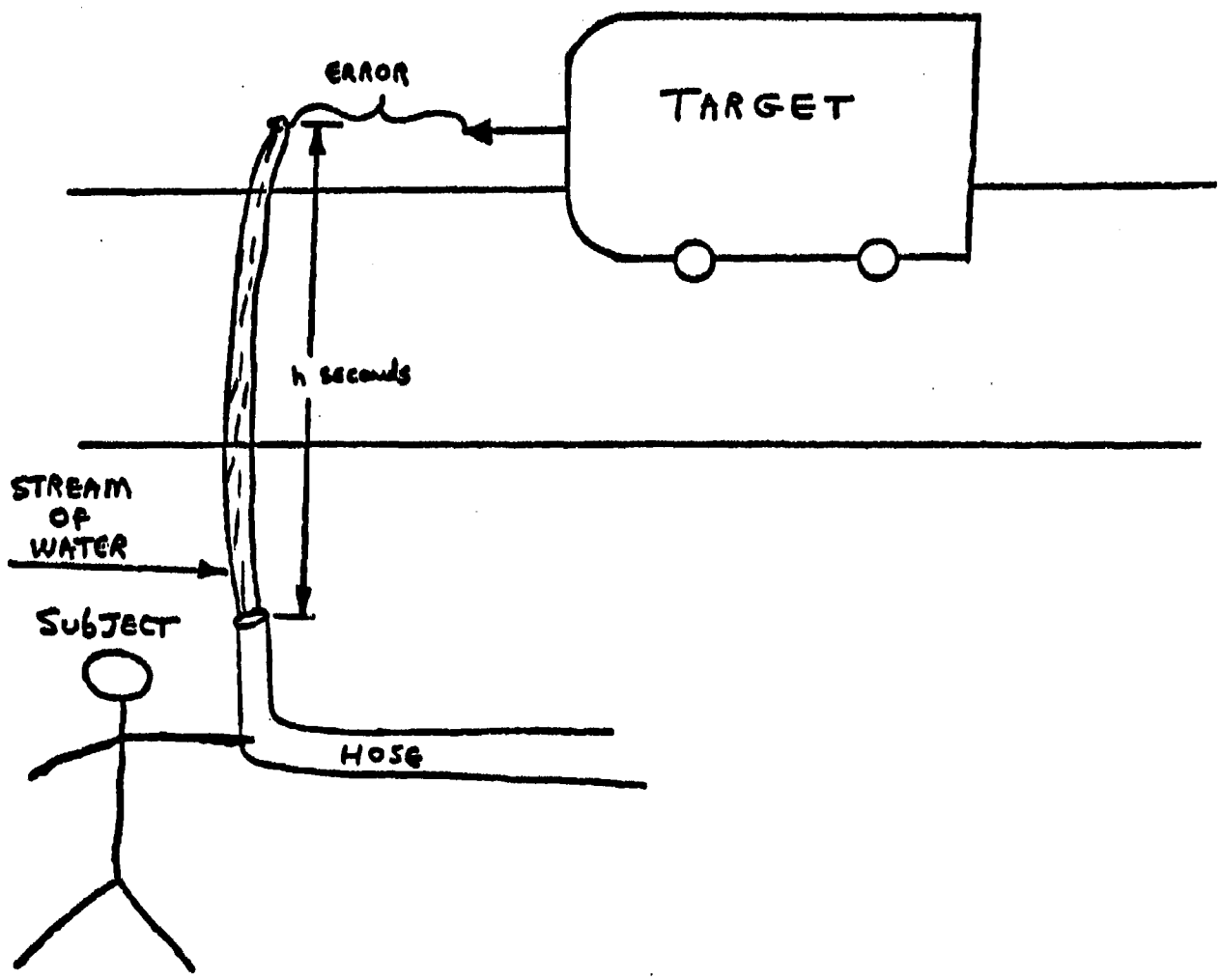


Figure (1) - A Tracking Task Analogous To The Tracer Fire System

A TRACER FIRE SYSTEM:

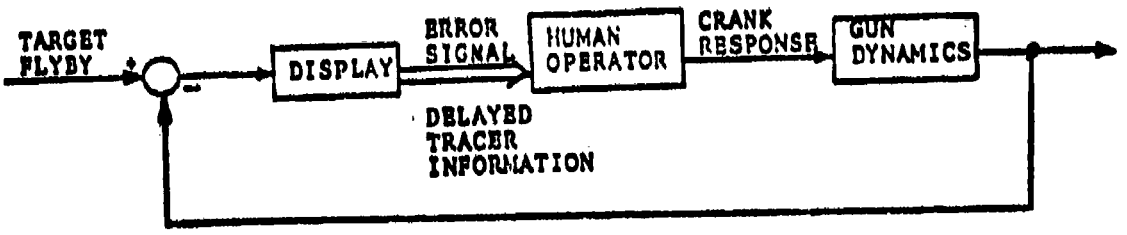


Figure (2a)

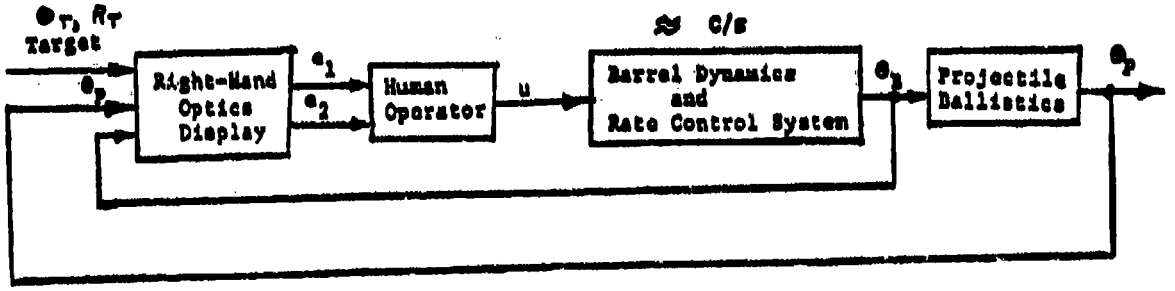


Figure (2b)

Display

Two Errors on The Display

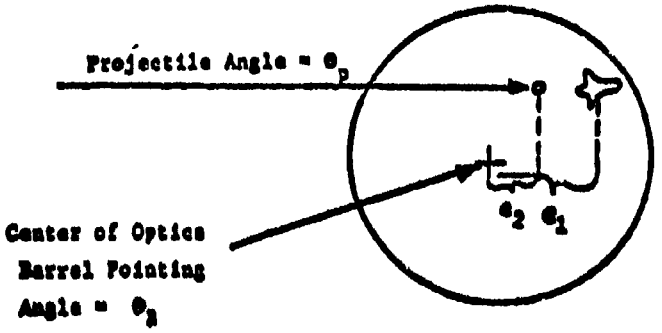


Figure (2c)

Figure (3a) - Azimuth Position

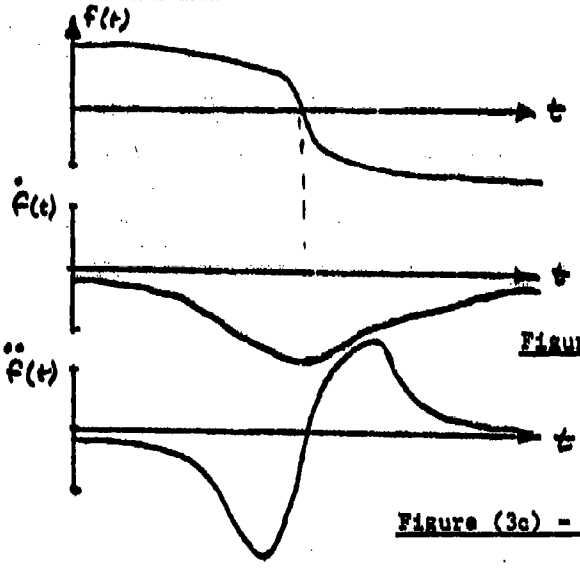


Figure (3b)- Azimuth Velocity

Figure (3c) - Azimuth Acceleration

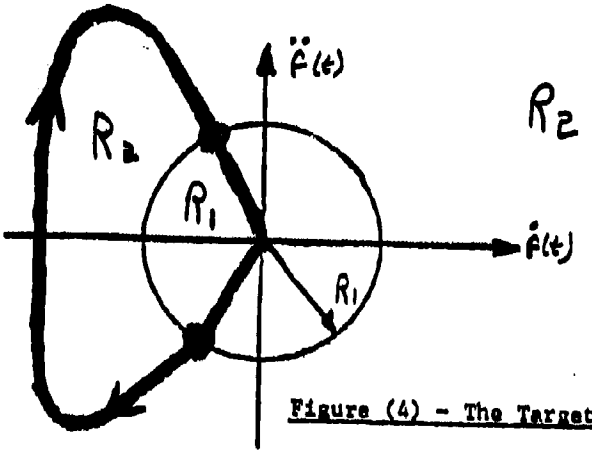


Figure (4) - The Target Phase Plane

Region R_1 - The Easy Tracking Region

$$\ddot{f}^2 + \dot{f}^2 < R_1$$

Region R_2 - The Difficult Tracking Region

$$\ddot{f}^2 + \dot{f}^2 > R_1$$

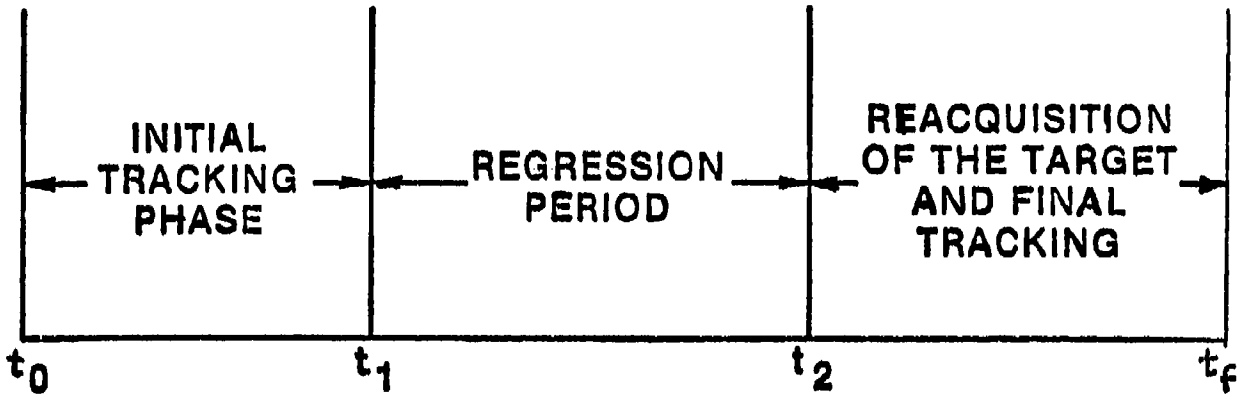


FIGURE (5) - A SEGMENTATION OF THE TRACKING TASK

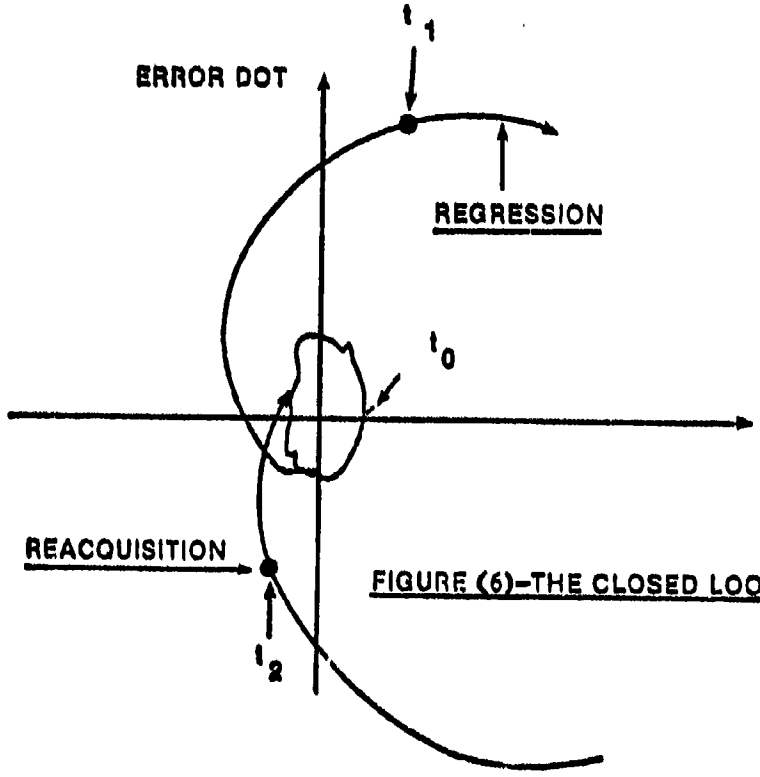


FIGURE (6)–THE CLOSED LOOP ERROR PHASE PLANE

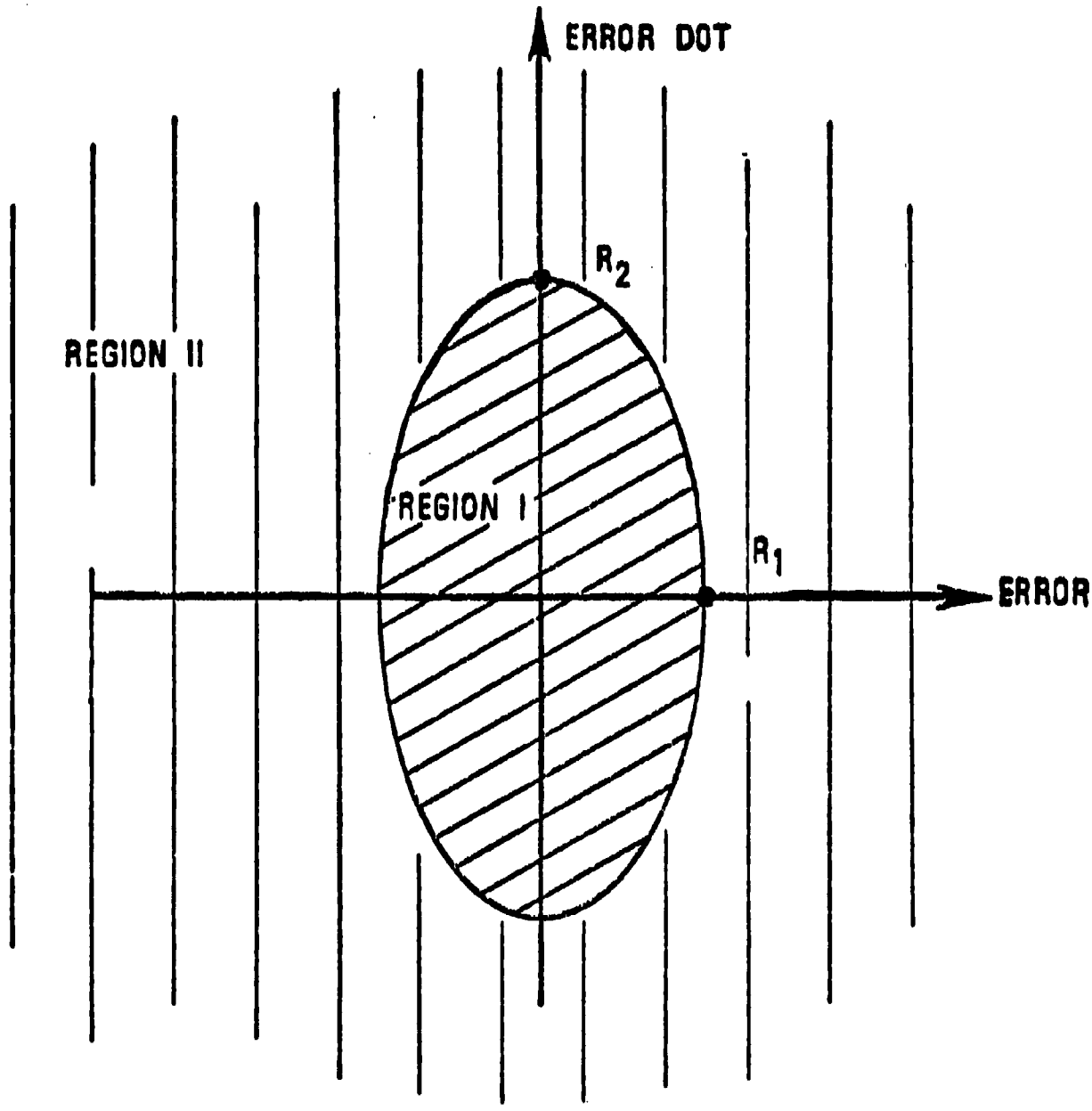
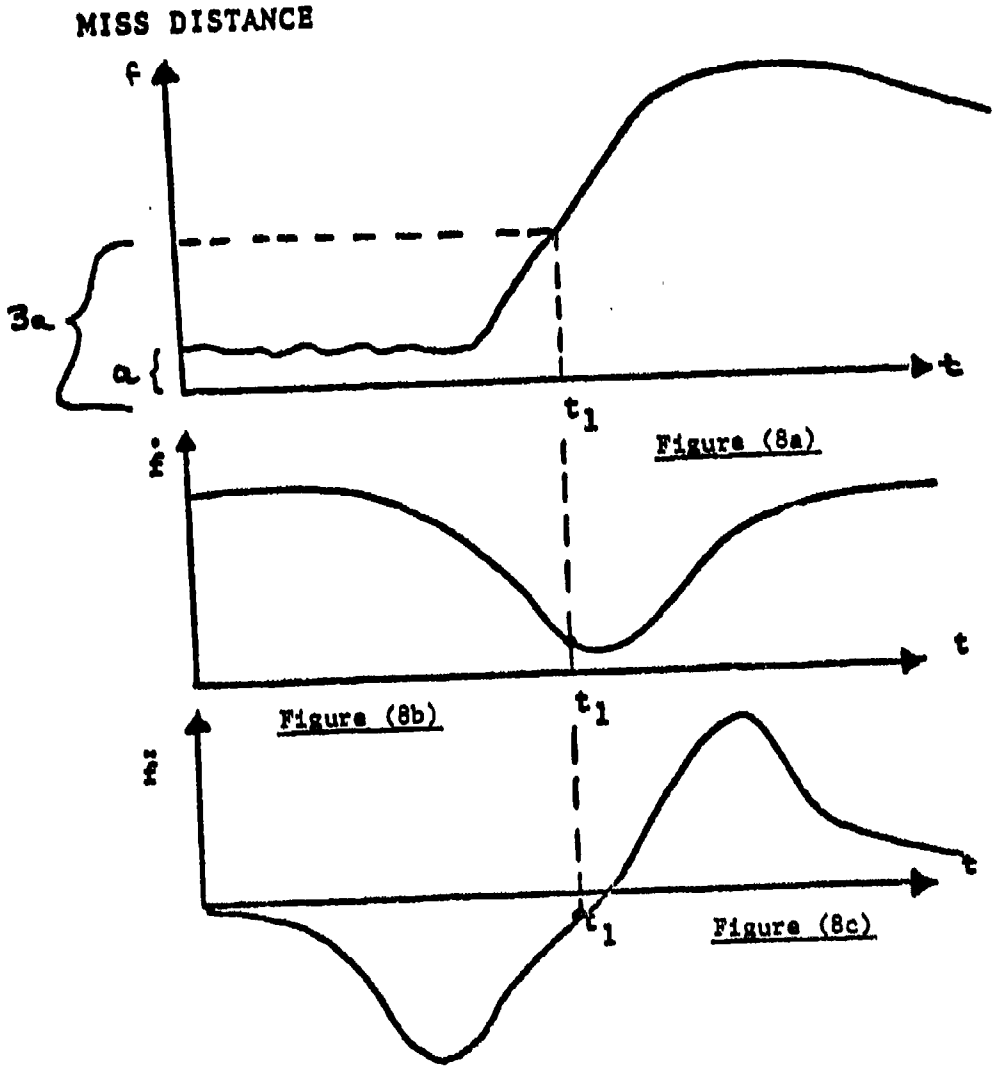


FIGURE 7 • THE BEHAVIOR REGIONS I AND II



Figures (8a-c) - A Heuristic Method To Determine Critical Boundaries

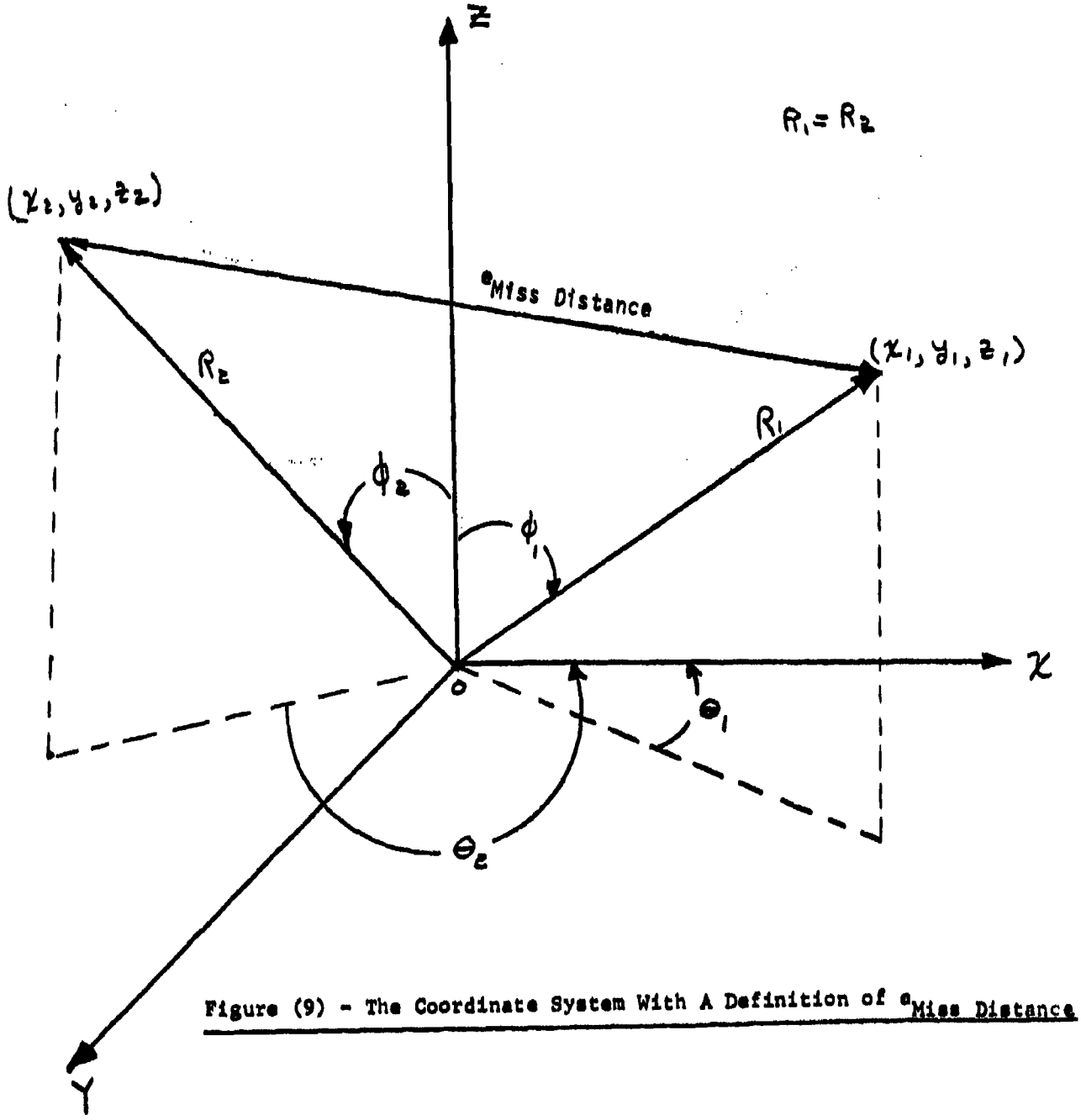


Figure (9) - The Coordinate System With A Definition of $e_{Miss\ Distance}$

Figure (10a) - Comparison of Trackers-Mode 6-Without Lead Angle Computer

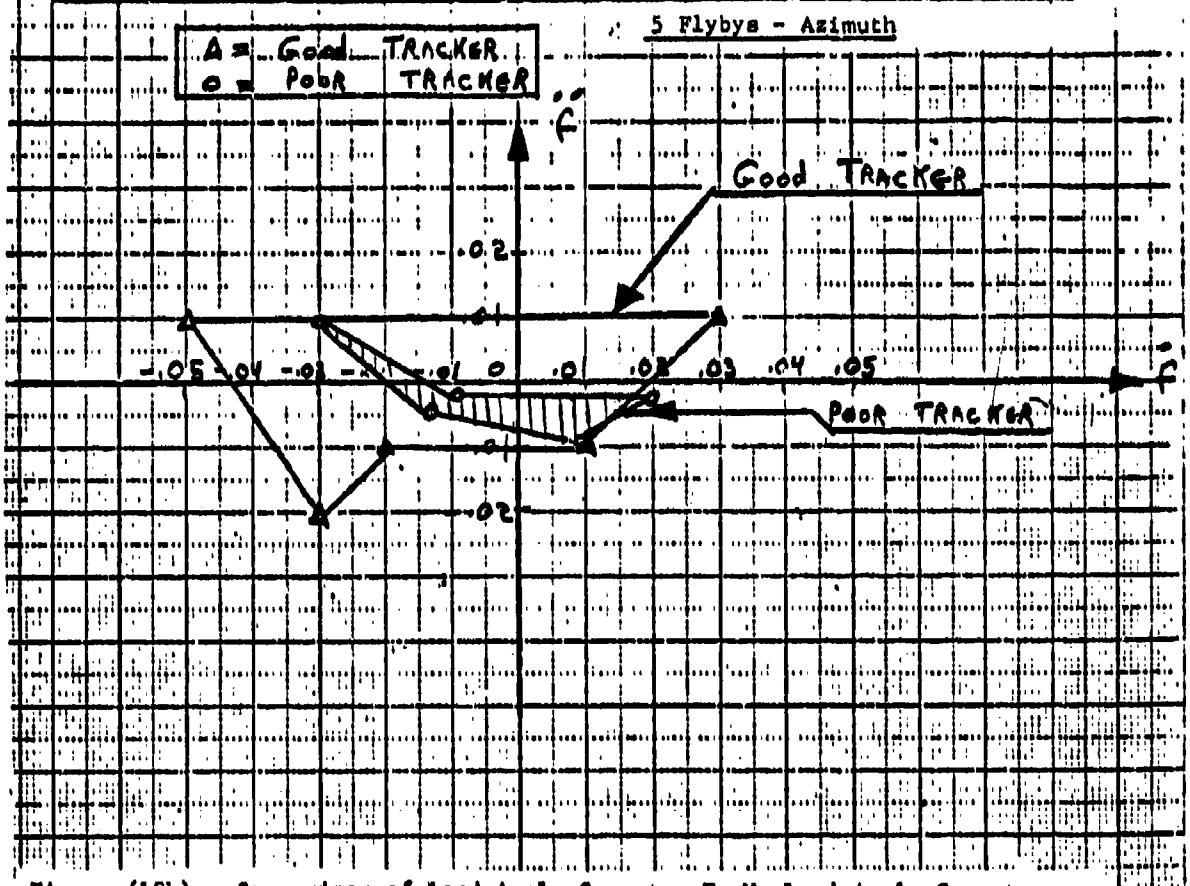
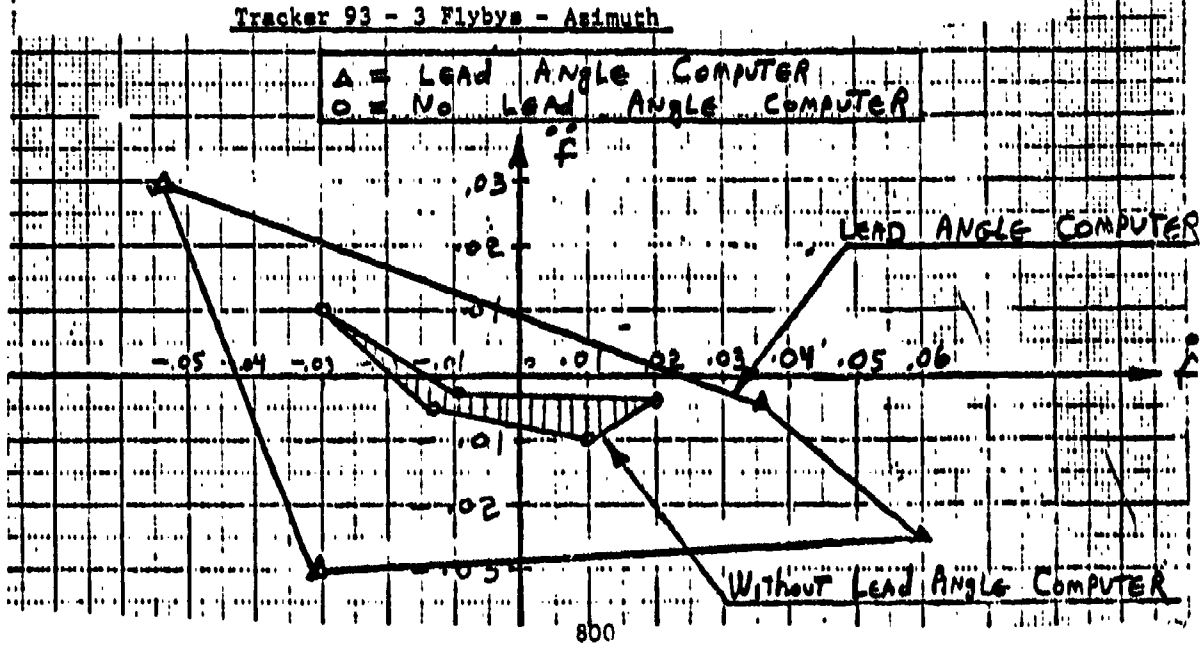


Figure (10b) - Comparison of Lead Angle Computer To No Lead Angle Computer



RESULTS FROM THIS ANALYSIS

Table I

MODE 2 - WITH LEAD ANGLE COMPUTER

SUBJECTS	AL	EL	AL	EL	AL	EL
	FB #1		FB #2		FB #3	
	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$
BEST-#33	-.05 .04	-.05 .003	-.3 -.14	.02 -.04		
MEDIOCRE-#80	-.05 .03	-.035 .002	-.1 .01	-.02 .001	-.13 -.001	-.01 .01
WORST-#33	-.05 .03	-.025 .001	.035 -.005	-.03 .001	.06 -.025	.005 -.001
	-.03 -.03	-.03 .001				

801

Table II

MODE 6 - WITHOUT LEAD ANGLE COMPUTER

SUBJECTS	AZ	EL	AL	EL	AL	EL	AL	EL	AL	EL
	FB #1		FB #2		FB #3		FB #4		FB #5	
	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$	$\dot{f}$ $\ddot{f}$
BEST-#80	-.02 -.01	.02 .002	-.05 .01	-.005 -.002	-.03 -.02	.04 .02	.03 .01	.001 .001	.01 -.01	-.01 -.01
MEDIOCRE-#33	-.01 -.01	.02 .001	-.04 .015	-.005 -.001	-.15 -.001	.002 .02	.025 -.001	.001 -.002	.01 -.01	-.01 .01
WORST-#93	-.01 -.001	-.001 .001	-.03 .01	-.005 .001	-.14 -.012	-.02 -.001	.02 -.002	-.001 -.01	.01 -.01	-.01 -.001

MANUAL CONTROL ASPECTS OF TACTICAL ANTIAIRCRAFT ARTILLERY AND SURFACE TO AIR MISSILE ENGAGEMENTS

George J. Valentino, Captain, USAF
6570th Aerospace Medical Research Lab (6570 AMRL/MEA)
Wright-Patterson AFB, Ohio 45433

ABSTRACT

The threat of AAA (Anti-Aircraft Artillery) and SAM (Surface-To-Air Missile) air defense systems against aircraft is known. In order to quantify this threat, many agencies are spending enormous resources on electronic warfare tests, survivability/vulnerability analyses, and tactics development. Efforts range from analytic paper studies through multi-player field tests. Unfortunately, many times only equipment performance is considered. At the Manned-Systems Effectiveness Division of AMRL, several programs are underway which address the biotechnology man-machine aspects of such engagements including human operator target detection, acquisition, tracking, fire control, and decision making. Man-in-the-loop laboratory experiments are used to quantify these sub-engagements; control theoretic human operator models are then developed for inclusion in Air Force standard computer engagement simulations. This paper will address the scope of these programs, techniques used, and results to date.

AUTHOR INDEX

LIST OF AUTHORS, 15th ANNUAL MANUAL CONFERENCE

Prof. Gyan C. Agarwal 277, 311
Gerry Albers 267
B.R. Ashworth 147
Sheldon Baron 3
Daniel Baty 650
Dr. A.K. Bejczy 404, 574
Denise C.R. Benal 301
R.A. Benel 301
Dennis B. Berringer 690
Hans E. Boller 457
Thurston L. Brooks 395
J.W. Browne 404
Keith Butler 71
B. Dell Campbell 669
Dr. Richard A. Chechile 71
Yee Yeen Chu 691
Warren F. Clement 555
Michael G.H. Coles 301
Bahman Daryarian 392
Dr. Ing Karl H. Doetsch 430
A.R. Ephrath 774
W.H. Faulkner 750
Richard D. Gilson 396
F.H. Glanz 43
Betty Glass 380
Daniel Gopher 93
Gerald L. Gottlieb 277, 311
T. Govindaraj 24
Joel Greenstein 31
William Gutowski 71
John E. Hart 445
Sandra G. Hart 632
Robert K. Heffley 545
Ronald A. Hess 369
R.J.A.W. Hosman 213
Richard J. Jagacinski 380
Robert J. Jaeger 311
Wayne F. Jewell 244
Henry R. Jex 244
Andrew M. Junker 168, 725
K.M. Kessler 515
David Kleinman 331, 487
G.D. Kontipidis 43
Jonathon Korn 331
L. Kretzschmar 300
Walter Krueger 457, 471
Shigeru Kurosu 388
Roy Lancraft 168, 487
William H. Levison 168, 187
J.L. Lewis 404
D.E. Limbert 43
Dr. Leonid Lipchin 268
Capt Terry L. Lutz 483
Ray Magdaleno 244
B.T. McKissick 147
Paul Milgram 524
R.A. Miller 114
Martin Moran 300
Ramal Muralidharan 3
David Navon 93
Heidi L. Neubauer 624
Sharon O'Connor 650
E.D. Onstott 750
Garrett Paine 574
Everett Palmer 71, 650, 703
R.V. Parrish 147
Anil V. Phatak 341, 515
Sheryl L. Prince (Chappell) 396
David L. Quam 243

Daniel W. Repperger 380, 725, 787
Dr. Ing Wolf Roeger 430
William Rouse 24, 31, 624
David K. Schmidt 749
John Senders 524
Thomas B. Sheridan 144, 392, 395
Richard S. Shirley 669
R. Steeb 691
Bernhard Tileman 502
Masayoshi Tomizuka 588
Pamela Tsang 82
George Valentino 802
G.A.J. van de Moesdijk 213
H.C. van der Vaart 213
Maris Vikmanis 787
Sharon Ward 380
Paulus Wewerinke 601
Christopher D. Wickens 82
Dana R. Yoerger 392
Greg Zacharias 187

*U.S.Government Printing Office: 1979 — 657-084/399